Bournemouth University

Virtual Technologies in Nurse Education

The Pairing of Critical Realism with Partial Least Squares Structural Equation Modelling as an Evaluation Methodology

Award of PhD

Heidi Singleton

18/09/2020

Abstract

Background:

Virtual technologies have been, and continue to be, of significant interest to Higher Education (HE) educators. There have been many research studies carried out into the efficacy and acceptability of these technologies. But, this research (via a systematic literature review) found that there are significant methodological shortcomings in many of those studies, particularly with respect to understanding the mechanisms of the effect of virtual technologies on learning. Most papers were superficial and concentrated on the Technology Acceptance Model (TAM) of usability and ease of use. Some carried out perfunctory assessments of learning effect, but predominantly by measuring student enjoyment via subjective self-reporting. This thesis responded to this gap in the literature by implementing a non-immersive virtual reality (VR) (accessed via a laptop), educational simulation of a deteriorating diabetic patient and creating a novel and powerful method to evaluate the effect of that simulation on nurse education.

Method:

The systematic review of the literature led to the creation of a diabetes VR simulator. A novel approach was designed to evaluate this simulator which consisted of the pairing of a randomised controlled trial (RCT) (n=171), analysed via Partial Least Squares-Structural Equation Modelling (PLS-SEM). The conceptual pathway model for this PLS-SEM approach was drawn from a Critical Realist (CR) review. Hence the main aim was to assess the effectiveness of CR paired with PLS-SEM as a method to evaluate the impact of VR simulations on undergraduate nurse education. The RCT enabled comparison of the VR simulation with normative teaching methods which addressed the two objectives: to determine the effect of pairing CR with PLS-SEM as an evaluative method, and to determine how using this novel evaluative method can inform our understanding of the impact and future use of VR simulations for undergraduate nurse education.

Findings:

The effect of pairing CR with PLS-SEM was that deep insight was gained into how VR simulations can benefit student nurses. VR was found to be significantly *(P=<.001)* better in terms of hypoglycaemia knowledge than normative methods. Moreover, the novel method also enabled identification of the key point of action of the simulation, via analysis of the conceptual model which evidenced that the “engagement to immersion” pathway was responsible for leading to higher knowledge scores in the VR group. This thesis is claiming addition to knowledge about how the novel methodological approach taken has the potential to deepen understanding of how virtual technologies can affect learning. Recommendations for policy, practice, and further research have been made on this basis. Future studies could use PLS-SEM combined with CR in order to ascertain both measurable and rich data about how new technologies can improve nurse education.

Acknowledgements

My Supervisors produced a research proposal that set the challenge of investigating scalability and sustainability issues relating to the use of Virtual Technologies for Higher Education. In response to this proposal I applied for a Vice-Chancellor's Scholarship, which became the initial funding for my Doctoral studies. Final year funding was provided by the Bournemouth University (BU) Faculty of Health and Social Sciences, when I joined as a member of staff in 2019.

First and foremost, I want to thank my three supervisors, Professor Debbie Holley, Professor Liz Falconer and Dr Jaqueline Priego. It has been an honour to be their PhD student. I appreciate all their contributions of time, ideas, feedback, coffee, meals, BBQs, and for making my experience productive, stimulating and enjoyable. The joy and enthusiasm they have for research was contagious and motivational. Each one has instilled in me research habits and approaches that will stay with me throughout my career. Professor Holley has always led by example with creativity and new ways of thinking. I will never forget the importance of "framing your writing" and the use of "punchy" abstracts and conclusions. Professor Falconer constantly inquired: "What are you trying to find out?" Her clear thinking and continued emphasis on focus and accuracy helped me to structure my work. Dr Priego also encouraged me to keep striving for excellence in all aspects of my research and shared her systematic review expertise with me. My supervisors have contributed immensely to my personal and professional time at BU. They have been a source of friendship as well as good advice and collaboration.

In addition to my supervisors, I am especially grateful for the collaborations and friendships with other teams. Firstly, I wish to thank Dr Janet James and Simone Penfold (Diabetic Nurse Specialist- Royal Bournemouth Hospital), who worked with me to create the deteriorating patient case study. Their expertise was invaluable. Dr Janet James also kindly allowed me to investigate the VR simulation within her teaching unit. Secondly, I wish to thank Daden Ltd. in particular, David Burden and Nash Mbaya. I have appreciated their collaboration and their impressive skills in turning the deteriorating patient case study into a prototype VR simulation. Funding for this was provided by the Centre for Fusion Learning Innovation and Excellence (FLIE).

Expertise was also sought from other BU colleagues. I extend many thanks to Dr John Beavis, who introduced me to the PLS-SEM approach, encouraged me to teach myself how to use SMART PLS software, and shared some of his expert statistical knowledge with me via numerous tutorials. In relation to Critical Realism, I am indebted to Professor Sam Porter for his engaging and innovative philosophical conversations. Thanks also goes to Dr Leslie Gelling who kindly agreed to read my thesis as a critical friend.

My Doctoral journey was also enriched by Dr Sue Eccles, Dr Jenny Hall, Dr Ann Luce, Dr David Biggins, Dr Curie Scott, Charlotte Webb and other members of FLIE who were always ready to help. I would also like to thank the Learning Technologists who taught me how to use the VR equipment and supported my fieldwork with students: David Hunt, John Moran, Stephen Pyne and Tracey Webb.

I am grateful for the assistance of the BU Library team including Emma Crowley (Search String and Review Protocol guidance), Caspian Dugdale (Endnote help) and Jan Hutt (assistance with referencing). Support was also provided by the IT Services team, who are all highly professional; Chris Pitt, was extremely helpful with formatting my thesis.

Thank you to the Doctoral College team and Faculty PGR (postgraduate research) support teams for providing advice and administrative assistance throughout. I am especially grateful for conversations with Professor Ian Donaldson and Professor Christos Gatzidis who provided initial and major review feedback, and to all staff members who volunteered to run PGR workshops, conferences and ethics committees.

My penultimate thanks go to my research participants, the second-year adult and mental health nursing students who volunteered to take part in this research, and Dr Sarah Eales who also agreed to be interviewed for this study.

Finally, I would like to thank my family and friends for all their love and for listening. Thank you to my parents, Anthony and Shirley, and my grandmother Rita, who raised me with a love of learning. My family supported me in all my pursuits, and many were medical corps and nursing role models. I am very grateful for all the meals, cups of tea and babysitting they provided as well. Thank you to my siblings, Ross and Aurelia for their encouragement. The understanding, love and faithful support of my husband Ben and young sons Ethan and Noah, during this PhD work is so appreciated.

Thank you to everyone mentioned and all those who helped during my Doctoral studies. Without them I could not have completed this thesis.

Heidi Singleton.

Bournemouth University October 2020

Author's Declaration

Some of the materials contained within this thesis have been presented at conferences as follows:

Date	Conference	Title	Authors
2020	NETNEP 2020 Oral presentation	Evaluation of a non-immersive virtual reality deteriorating patient simulation in higher education	Singleton, H. James, J. Beavis, J. Mbaya, N. Falconer, L. Holley, D. Priego-Hernandez, J. Penfold, S.
2020	Diabetes UK Professional Conference 2020 Poster	Virtual reality can improve nursing students' knowledge of hypoglycaemia; do students prefer immersive virtual reality to non-immersive virtual reality?	James, J. Singleton, H. Penfold, S. Holley, D.
2019	Diabetes UK Professional Conference 2019 Poster	Assessing student nurses' knowledge of hypoglycaemia- the first step to improvement.	Singleton, H. James, J. Penfold, S. Priego-Hernandez, J.
2019	Café Scientifique Oral Presentation	A 'wicked challenge': supporting our students' learning with new technologies	Singleton, H. Holley, D. Falconer, L.
2018	UCISA Conference Oral presentation	Engage and Educate with Virtual Reality: sharing our experiences of VR Deteriorating Patient Oral presentation	Singleton, H. Holley, D. Falconer, L.
2018	OER18 Bristol Oral presentation	Virtual Reality: the implications for open educational resources Oral presentation	Singleton, H. Holley, D. Falconer, L. King, D.
2018	BU PGR Conference Poster	Evaluation of the effectiveness of virtual, augmented and mixed reality technologies for learning in higher education: a systematic review of the literature	Singleton, H.

The Line of Argument within this Thesis	Location within the thesis
Virtual technologies have been, and continue to be, of significant interest to HE educators. There have been many research studies carried out into the efficacy and acceptability of these technologies.	1.0 Introduction
But, this research (via a systematic literature review) found that there are significant methodological shortcomings in many of those studies, particularly with respect to understanding the mechanisms of the effect of virtual technologies on learning. Most papers are superficial and concentrate on the Technology Acceptance Model (TAM) of usability and ease of use. Some carried out perfunctory assessments of learning effect, but predominantly by measuring student enjoyment via subjective self-reporting.	2.0 Systematic review
This research has identified Critical Realism coupled with PLS-SEM as a methodological approach that could enable us to understand the pathways to learning that students experience when using virtual technologies, thereby adding to our knowledge about both virtual technologies in education, and about the methodological approaches that might move the evaluation field on, past its current superficial approach.	4.0 Method
The Critical Realist (CR) literature review identified the latent variables that form the pathways that were tested in the experiments. Therefore, two literature reviews were conducted.	3.0 Critical Realist Review
The experiments were in the field of health practitioner education - specifically nurse education in respect of diabetes management. Hence, the thesis is not claiming generalisability from the study, but addition to knowledge about how the novel methodological approach taken has the potential to deepen our understanding of how virtual technologies might affect learning. Recommendations for policy, practice, and further research have been made on this basis.	7.0 Discussion and 8.0 Conclusion

Contents

List of Figures

List of Tables

Glossary of Frequently used Technical Terms

Technical Term	Abbreviation Used
Augmented Reality	AR
Bournemouth University	BU
Contextual Mechanisms + Programme Mechanisms + Agency = Outcome	CPMAO
Critical Realist/Critical Realism	CR
Heterotrait-Monotrait Ratio	HTMT
Higher Education	HE
Learning Technologist	LT
Measurement Invariance of Composite Models Procedure	MICOM
Mixed Reality	MR
Partial Least Squares-Structural Equation Modelling	PLS-SEM
Randomised Controlled Trial	RCT
Second Life	SL
Technology Enhanced Learning	TEL
Variance inflation factor	VIF
Virtual Reality	VR
Virtual Reality Environment	VRE
Virtual World	VW

Chapter 1 **Introduction**

1.1 Synopsis

"A mind that is stretched by a new experience can never go back to its old dimensions" (Holmes 1841-1935).

Nurse education is continually challenged by the “theory-practice gap” (King et al., 2018), notably for complex clinical procedures and clinical decision-making (Weeks et al., 2019, El Hussein et al., 2019). Nurses have reported difficulties understanding the bioscience underpinning nursing care (Branney and Priego-Hernández, 2018). Educators are constantly faced with finding accessible and innovative methods of teaching and learning that transition the student from novice practitioner to work ready clinician (Butt et al., 2018, Irwin and Coutts, 2015). Responding to this, a virtual reality (VR) deteriorating patient simulator was designed and evaluated, for use with adult and mental health nursing students, to provide direction for spanning this “theory-practice gap” in the specific area of hypoglycaemia treatment.

A novel evaluative approach was developed to enable greater understanding of the pathways to learning that students experience when using VR technologies, thereby adding to our knowledge both about VR technologies in education, and the methodological approaches that might move the evaluation field on, past its current superficial approach.

In this first chapter, the background to the thesis is set out, including: a discussion of the technical terminology used, the affordances and barriers to the adoption of Virtual and Augmented Reality use in education, a summary of pedagogical methods used in Higher Education (HE), a basic introduction to hypoglycaemia and a summary of the use of simulation in nurse education.

1.2 Rationale for conducting this research

This research will bridge some of the gaps in effective evaluation of VR by implementing original, empirical research exploring theoretical instructional design and evaluation of these technologies for the student body, with a plan for scalability across contexts. It has a strong rationale in view of the rapid evolution of VR technologies. The appraisal of such technologies has been somewhat superficial and has relied upon reports of student enjoyment and acceptance. Immersive Virtual Worlds (VWs) such as Second Life (SL) have been deployed in HE for more than a decade, with research indicating strong student engagement (Philips et al., 2015), but often relying on self-reporting for learning outcomes (Kurilovas et al., 2016). Augmented Reality (AR) applications, in turn, are task-specific tools that integrate digital information with the user's environment; they have wide applications but, again, the evaluation of their impact upon learning has been limited (Martin-Gonzalez et al., 2016, Martin-Gutierrez et al., 2014).

As pioneering universities around the world gradually shift from conventional multimedia to more immersive and interactive VR and AR technologies (Newman, 2020), the question of effectual pedagogy comes to the forefront. Educationalists endeavour to find guidelines that could assist them in developing and adopting original innovative learning systems. Technologies such as VR and AR could bring an abundance of learning resources to Educational Institutions across the world, but how, exactly, can they be blended appropriately, and how can they be implemented in an authentic manner?

VR technology has not yet become embedded into curricula; and whilst the challenges associated with its adoption have been reviewed (Akcayir and Akcayir, 2017) there remains a lack of clarity concerning the underpinning theories and instructional design aspects (Merchant et al., 2014). Steils et al.

(2015) believe that change has resulted in curricula with relatively little pedagogical foundations and a trend towards technological determinism. They called for a mapping of VR theories, pedagogies and practices in order to delineate their impact; this call will be responded to by the systematic review of the literature. This research seeks to contribute to bridging this gap in the knowledge, via systematic review, and contribute to pedagogical theories aligned to VR use in HE.

In this thesis, the relationship between innovative pedagogical practices using technology enhanced learning (TEL) will be explored and prospects of sustainability, scalability and transferability of these innovative practices in HE settings will be proposed. The main obstacles faced by academics are time and funding. Historically, funding is an ongoing problem for TEL, especially as rapid changes outstrip the opportunities to evaluate tools in depth (Bakir, 2016). A Realist approach will be applied throughout, involving: systematically appraising the literature in respect of the two themes; operationalising learning innovation by developing a learning and teaching approach (using VR and other tools for scalable immersive experiences; e.g., mobile devices); and evaluating the learning potential of the approach.

Most studies in the field of VR have only concentrated on a comparison of desktop VR compared to PowerPoint slideshows (Parong and Mayer, 2018), or video. The research reported in this thesis compared VR learning against a paper case study; the latter was the “normative instruction”. A search of the literature revealed that considerable research, until now, has been descriptive in nature and the methodological quality of the few quantitative studies was variable (Allcoat and von Mühlenen, 2018, Ogbuanya and Onele, 2018). Where randomised controlled trials (RCTs) have been conducted sample, sizes have not always been appropriate. The study presented in this thesis incorporated a larger scale RCT (n=171).

1.3 Background to VW/VR/AR

In this thesis the term ‘virtual technologies’ refers to the mix of devices and platforms that enable users to experience and/or interact with simulated environments and artefacts in real time; specifically, virtual worlds (VWs), virtual reality (VR) and augmented reality (AR). Passive watching of 3D video has not been included in the study, as the lack of interactive capability makes this a substantially different educational experience. Table 1 summarises the definitions, affordances, and challenges of the virtual technologies that are the subject of this thesis.

Table 1- Definitions, learning affordances and challenges of VR/VW/AR technologies.

VIRTUAL WORLDS	
Definition	VWs are shared, simulated and persistent spaces which are inhabited and shaped by their inhabitants through the agency of avatars. These avatars mediate our experience of this space as we move, interact with objects and interact with others, with whom we construct a shared understanding of the world at that time (Amin et al., 2016)
Key Learning Affordances	A fully immersive experience that gives the user a strong sense of being there (Warburton, 2009, Morales et al., 2019) Collaborative tasks are possible even for distance leaners (Falconer, 2017) With inbuilt construction and programming tools and without the boundaries of gaming environments, learners can engage in a process of trying and testing, rather than following a pre-existing design (Dreher et al., 2009)
Challenges	There is a need to set rules for appropriate in-world behaviour (Ward et al., 2016) Different time zones if users are international (Kawulich and D'Alba, 2018) There may be an upfront investment in time and learning (Ward et al., 2016, Wang and Burton, 2013)

VIRTUAL REALITY

Definition	VR consists of 3D, fully immersive headsets with 3D sound, and can also include haptic devices. It is a set of technologies that enable people to immersively experience a world beyond reality (Berg and Vance, 2017) VR displaces a person to another location, with complete immersion as the goal (Brigham, 2017)
Key Learning Affordances	Users can work in controlled, safe, cheap, artificial spaces (Rogers, 2011, Merchant et al., 2014) Users can access (and re-access), learning from any location at any time (Falconer, 2017, Minocha and Reeves, 2010, Penland et al., 2019) Enables embodied cultural interaction, such as bowing in Japanese greetings (Cheng and Tsai, 2013)
Challenges	User disorientation, nausea and headaches (Barrett, 2004) Cost concerns (Cruz-Neira et al., 2018, Elkoubaiti and Mrabet, 2019) There is a learning curve to understanding the interaction of the system (Raikwar et al., 2019)

AUGMENTED REALITY

Definition	AR is a type of VR in which synthetic stimuli are superimposed on real-world objects usually to make information that is otherwise imperceptible to human senses perceptible (Lopreiato et al., 2016) Users can overlay virtual graphics or audio (Azuma, 1997, Behmke et al., 2019)
Key Affordances	Capacity to access outside resources to problem solve (Klopfer and Squire, 2008) Physical world annotation, contextual visualisation and vision haptic visualisation (Bacca et al., 2014, Singhal et al., 2012) Digital storytelling and other interactive activities (Marcel, 2019, Godwin-Jones, 2016)
Challenges	Users must be in the physical environment to which the AR relates (Oleksy and Wnuk, 2017) Educators must integrate and manage the AR experience (Dunleavy et al., 2009) Pedagogical issues, e.g., a need for more class time (Akcayir and Akcayir, 2017, Kamarainen et al., 2013)

The relationship between AR and VW/VR is a relatively unexplored area, along with the nature of the boundary between virtual and physical, and how transitions across those boundaries may influence learning (Savin-Baden and Falconer, 2016). AR and VR are frequently mentioned in the same breath, as they share many features; however, there are significant yet blurred distinctions between the two. In the literature, there are also many variations in terminology associated with VR; this has led to confusion over definitions. To address this, the following paragraphs aim to clarify the definitions of these technologies and their associated affordances and challenges.

As Brigham (2017) pointed out, fundamental elements of VR were proposed and used as early as the 1940s. Following early flight simulation efforts by Richard Day, and NASA's "Virtual Visual Environmental Display" developed in the early 1980s; the term Virtual Reality was created in 1989 by Jaron Lanier (Kelly and Lanier, 1989). He stated that VR:

> *"Recreates our relationship with the physical world in a new plane, no more, no less. It doesn't affect the subjective world; it doesn't have anything to do directly with what's going on inside your brain. It only has to do with what your sense organs perceive."*

This perception is outdated, as the way VR is viewed has changed over the years. VR is no longer linked to what is "going on inside your brain". Current understandings of how VR works are more concerned with issues such as embodiment, agency, presence and immersive tendency. These terms are frequently used when discussing VR and will be drilled down further within Chapter 3 (Critical Realist Review of the Literature). However, a brief overview is presented here.

1.3.1 VR Key Concepts

Starting with embodiment and agency, researchers have long considered how best to describe the way we represent and experience ourselves. In order to move towards a definition, they have separated out the concept into three main categories: sense of agency—the experience of causing world actions/events (Ellis and Newton, 2005) sense of embodiment—the experience of owning a body and identifying with the location of that body (Longo et al., 2008); and sense of presence—the experience of being “situated” in an environment (Perez-Marcos et al., 2012).

One part of how we perceive ourselves is through agency, which has sometimes been viewed as part of embodiment and on other occasions has been investigated as an independent construct. For example, in one line of research, hypnotic suggestions have been shown to encourage changes in the way that predisposed participants generate/monitor actions, leading to alterations to their sense of agency (Polito et al., 2014). Sense of agency has been measured in a variety of ways including explicit ratings of first-person experience and indirect, implicit measures such as intentional binding, which uses participants' time judgements regarding causal actions (Longo et al., 2008) in a behavioural task as a proxy for agency.

Another way of perceiving how we “experience” ourselves is via embodiment. Users who experience a strong sense of embodiment in their avatars tend to perceive avatar actions as their own. In addition to embodiment, people can understand and empathise more when they comprehend another person's experiences. VR content can stimulate empathy towards others (Dyer et al., 2018). VR can convey another person's experience or feelings to a viewer. In VR environments, viewers may strongly feel another person's emotions or situation by being in the same space, close to that character. Furthermore, stimulated empathy in VR can increase a user's overall empathy and the perception that a virtual environment is realistic. Through this empathy, users can feel a sense of embodiment or embodied cognition based on the stories (Pamungkas and Ward, 2015). In other words, the embodied cognition in VR helps users feel a sense of embodiment (Hofer et al., 2017). Fully immersive VR offers a sense of embodiment, in which users see themselves as part of the VR environment.

Presence is defined as “the subjective experience of being in one place or environment, even when one is physically situated in another” and is sometimes referred to as mental immersion (Tüzün and Özdinç, 2016, Hoffmann et al., 2006, Farrow and Iacovides, 2014). Users with a strong sense of presence observe the virtual environment as the reality that enfolds them rather than as images on screens (Slater, 2018). In several studies, operators have the discernment of being a fragment of the virtual environment (Slater, 2018). The notion of presence can be separated into detachment from and immersion in the virtual environment.

Presence can be influenced by regulator features, corporeal factors, disturbance factors and realism factors (Witmer and Singer, 1998). The minimisation of distractors intensifies the preoccupation of the user in the environment and positively influences their contribution. Writings on presence propose that the interstices amid what is physical and what can be computer produced is becoming progressively indistinct within technology (Schubert et al., 2001). From the cognitive point of view, the more the user is consciously attentive (Riva, 2008) the greater their sense of presence in a VW.

From a technological angle, presence is a response to immersion, defined as a technology’s ability to produce a credible, engaging environment with which the user can interact (Slater, 2018). Consequently, presence is improved by decreasing lag in the display and enhancing the fidelity of the projected scenes. This is something that could be challenging to achieve with cheaper VR technologies. According to Winn and Jackson (1999) when we lengthen our arena of sight onto a computational environment past about 60 degrees, an extraordinary phenomenon happens. We transfer from a sensation of observing a picture to a sensation of being in a place. We transfer from being spectators to participants, from looking at a display to "inhabiting an environment". In a similar vein, Kant, cited in Rescher (2000) elaborated on how we can never completely distinguish reality past the array of sensibility, how "knowledge cannot extend its domain over everything which the understanding thinks". The Kantian perspective is that past what we intentionally observe, outside the limits of experience, we have no theoretical structure to distinguish the thing. What we can tell is conditioned by our experience. Grasping a sense of existence in VR induces opposing these concerns.

Moving now to immersion, Witmer and Singer (1998) view immersion as the psychological response to the technology and describe it as a "psychological state" and state that the "degree to which they feel immersed in the VE" is by successfully separating users from the physical world. In contrast to this viewpoint, Slater and Wilbur (1997) define immersion as a technical feature of a VR system and define presence as a result of an immersive technology. For Witmer and Singer (1998), immersion involves internal factors of the individual, and is:

> *"a psychological state characterized by perceiving oneself to be enveloped by, included in, and interacting with an environment that provides a continuous stream of stimuli and experiences."*

This definition is very similar to Slater and Wilbur's (2018) use of presence. As is evident, many of these terms are closely intertwined and associated.

The definition adopted in this thesis is that immersion is the degree to which a technology can achieve an inclusive, extensive, surrounding and vivid illusion of reality for the users' senses, matches the users' movements to the visualisations of the IVR, and presents a convincing plot to the senses of the user (Hauze et al., 2019). Even if immersion seems to be a crucial element for VR, as Freina and Ott (2015):2 say, VR can also be non-immersive when it:

> *"places the user in a 3D environment that can be directly manipulated, but it does so with a conventional graphics workstation using a monitor, a keyboard, and a mouse."*

It is this non-immersive VR that has been evaluated by research reported in this thesis.

To bring the discussion surrounding the underlying VR principles to a close, immersive tendencies will now be briefly defined and explored. The term immersive tendency is prevalent in the literature (Yi Fei et al., 2017, Dang et al., 2009). People who have more powerful immersive tendencies will report a greater feeling of presence in virtual environments (Riley and Kaber, 1999). Newman (2020) defines people with elevated immersive tendencies as individuals who:

> *"are able to block external distractions and become very focused, to the point where they become unaware of their immediate environment and the passage of time."*

Immersive tendencies are "thought to be dependent on aspects of human cognition and behaviour, including concentration, imagination, and self-control" (Riley and Kaber, 1999). Further investigators have established a connection between daydreaming and becoming lost in books and immersive tendencies (Falconer et al., 2018). The notion of immersive tendencies will be revisited throughout the rest of this thesis.

Having identified some key concepts that underpin VR, the narrative now turns to the progress of VR developments over the last few decades. Though some foretold AR and VR would be mainstream by the conclusion of the 20th century, it has taken time for corporations to be involved in financing such a technology and for the value to fall to an acceptable level (Brigham, 2017). Former VR systems (circa 1990s), were not only costly but furthermore had the challenge of making operators nauseous; two challenges which did not improve for another 20 years.

Latency, in the instance of VR, is the time between when somebody moves their head and when the virtual image corrects. Too much latency in VR leads to an operator's visual system to diverge with the vestibular system, and produces motion sickness (Attaran and Morfn-Manibo, 2018). In 2014, Palmer Luckey (founder of the Oculus Rift™) discovered how to develop a device that reduced latency. Gyroscope, accelerometer, and magnetometers combine to assess head motion, and consequently improve the operator practice. The accomplishment of the Oculus Rift™ encouraged a new curiosity in VR technology. Additionally, success of AR-based Pokémon GO, Harry Potter AR and developments from numerous corporations such as Facebook, Samsung and Google, indicate the start of a potential proliferation surrounding AR usage. The technology will now be discussed starting with VR.

1.3.2 Virtual Reality and Virtual World Definitions, Affordances and Challenges

Typically, VR technologies can be defined by the following features: they are inhabited by avatars that characterise users; they have several users; they deliver the illusion of movement in 3D space; and they have interactive chat functions (Radford et al., 2011). VR utilises a computer to generate a simulated environment. Users are inside that replicated world, rather than externally looking in, as in AR. Wider VR adoption has been made possible through advances in technology including processing speed and multimedia capability.

VR technologies usually consist of 3D, fully immersive headsets with sound, and can also include haptic devices that are worn on various parts of the body. According to Brigham (2017), VR completely displaces a person to another location, with complete sensory immersion as the goal. It conceals the user’s physical surroundings through use of head-mounted displays (HMDs) and replaces them with a computer-generated scene.

A considerable amount of literature has been published on VR and it has become a niche in certain fields, notably in the military, engineering and architecture (Samsudin et al., 2014). For instance, Dunleavy et al. (2009), described how engineering students tasked with virtual assembling of an object, completed the task more efficiently when using an immersive headset (Oculus Rift^{TM}) and haptic gloves.

Despite such advantages, VR systems can be associated with high cost because specialist hardware and high-level graphics computing capability are needed. Another challenge when using VR is reported user disorientation, nausea and headaches, caused by anthropomorphism, that leads us to believe that the experience is ‘physical’ as opposed to virtual (Barrett, 2004). Spatial constraint is a further challenge of VR; specifically, there are dangers associated with moving around in an uncontrolled environment when your senses of sight and hearing are compromised. Users wear a headset and headphones that providing surround-sound effects, all powered by a high-performance computer/gaming system or phone. Some VR systems come with laser sensors positioned around the room, and controllers that allow users to interact with the VW.

VWs can be viewed as a subset of VR technologies. Bell (2008) defined VWs as persistent, computer-generated environments with which users interact through the agency of avatars. These environments can be accessed either as fully immersive VR experiences, or on computer device screens. VR technologies are an immersive way of accessing a virtual environment, but such environments can also be accessed without VR equipment. Hence, user experience of virtual environments can differ, depending upon the technologies employed to interact with them.

The absolute definition of VWs remains disputed (Warburton, 2009). The rapid evolution of virtual technologies continues to expand and confuse definitions of VWs, as well as VR and AR. The copious contextual explanations that have surfaced, from the viewpoints of authors, researchers, industry specialists and the media, have added complex agreement on a shared understanding of VWs. Warburton (2009) did, however, note that a VW delivers a practice set within a technological setting that provides the user with a sturdy feeling of “being there”.

Savin-Baden and Falconer (2016) retained Bell’s (2008) ‘persistent’ element of VWs in their five identified characteristics:

- Persistent (changes made in the world remain)
- Synchronous (participants in the world are present at the same time, regardless of location in the world)
- Social (they allow for user interaction)
- Visually interactive (users interact using avatars)
- Visually rich (containing a wide range of detailed 3D environments)

These “social” and “synchronous” elements have remained a focus of Falconer’s (2013) research because they are advantageous to contemporary, HE teaching and learning approaches.

There have been a rising number of studies exploring the outcome of VWs on learners' education across diverse disciplines such as language, health care, architecture, business, and literacy (Bower et al., 2017). The findings of these studies have shown that VWs can be used as an environment in which learners are involved in education experiences. Dalgarno and Lee (2010), emphasise that technologies themselves do not directly enhance education but can afford particular learning tasks that themselves may result in learning benefits. SL is one such virtual online world popularly used in HE (Ata, 2016) which allows individuals to exist and engage with each other through assumed online identities (called avatars) and participate in pedagogic activities.

A key advantage of VWs is that they require students to be an active learner (Inman et al., 2010, Kirriemuir, 2008). Moreover, not only do VWs promote active learning, but they can also provide experiences that learners would not usually have. For example, Rogers (2011) found that Nursing students were able to work in an artificial space performing tasks that would be too costly, unsafe and risky to perform in the real world.

Collaboration has been highlighted, as an advantage of VWs, by several researchers (e.g., Falconer, 2013), with Dalgarno and Lee (2010) noting that:

> *'three-dimensional virtual environments that allow learners to engage simultaneously in shared tasks and/or produce joint artefacts by operating on the same objects in real time can pave the way for rich and truly collaborative experiences that foster positive interdependence within a learning group' (ibid. 2010:22).*

Falconer (2013) believes that a collaborative feature of VWs is not only a significant educational affordance but that it is also frequently overlooked.

She stipulated that VWs permit students to conduct collaborative tasks amongst groups who are physically long distances apart, or who are disabled and unable to take part in collaborative learning experiences physically. This insight was gained through many years of teaching a Master's unit of work, entirely through VWs involving students around the world. It is now generally agreed, from a range of studies, that VR (including VWs) is a stimulating and enjoyable way to learn (Warburton, 2009, Makransky et al., 2019, Nussli et al., 2014). This study will also explore if it also impacts learning.

Moving on from VWs, and returning to the general term VR, VR has become a niche in certain fields, notably in the military, architecture and engineering (Samsudin et al., 2014). There has been research conducted to investigate the use of VR in health care. The 2016 Horizon Report provided examples of how Nursing Students use Oculus Rift™ headsets to absorb techniques for catheterisation using VR. The technology set-up is less costly and necessitates less space than traditional medical mannequins, as well as permitting real-time feedback. Student Nurses at Kingston University and St George's University (UK), learn in a Paramedic Clinical immersive VR suite; which allows trainees to practise giving medical care while undergoing the difficulties and volatility of hectic atmospheres such as roadside accidents. Learners described that use of the amenities caused amplified confidence and improved communication abilities (Lindgren and Johnson-Glenberg, 2013).

Despite these advantages, conventional immersive VR systems are costly, fragile and not appropriate for lengthy periods of use (Limniou et al., 2008). Certainly, running VR systems is linked with high cost since special hardware is required, such as HMDs or numerous projectors, and 3D input devices. These disadvantages make immersive HMDs problematic to use in large classroom situations.

Another challenge when using VR is reported user dizziness. Virtual reality is chiefly about convincing the brain (anthropomorphising) that a computer-generated 3D environment conveyed to your eyes via a headset is 'real', a concept known as 'presence’. Anything that does not look 'right' to our visual system may deliver a sub-par experience, or even cause disorientation, nausea and headaches (Pettijohn et al., 2019).This is less of a problem with AR, because you can still see the real world.

Evaluation of the efficacy of VR as an approach to enhance student learning has been inconclusive thus far. Whilst Goeser et al. (2011) stated that research has demonstrated improved engagement when learning with VR, Fernandez (2017) highlighted “barriers” to using VR technology for educational purposes. Fernandez (2017) believes that such barriers include educators, faculties and institutions. Often educators might feel that they do not have enough time, training or resources to provide TEL opportunities for their students. Funding for such innovations might not be in line with faculty or institution aims or objectives. However, now that low cost, easy to use tools, including Google Cardboard and 360 cameras have become available, it is hoped that such exciting innovations will proliferate. Quality research will also provide confidence to managers and educators that such interventions are effective and desirable within HE teaching settings.

1.3.3 Augmented Reality Definitions, Affordances and Challenges

Having defined VR technology, AR will now be described and defined. AR is sometimes referred to as Mixed Reality (MR); however, upon closer inspection, such activities can be classified as either AR or VR. Indeed, Freina and Ott (2015), in their systematic review, use the search string: “augmented OR mixed reality” hence highlighting their interchangeable nature. Therefore, whilst it is acknowledged that researchers are describing their technology as MR or hybrid, for the remainder of this thesis AR or VW/VR or both will be referred to.

The term AR was first used to describe technology by Tom Caudell in 1990 while developing ways to visualise constituent information in aircraft design. Following prototypes that were developed during the 1990s, the first AR system was established by Louis Rosenberg in 1992. According to Holley and Hobbs (2020), the defining feature of AR is that:

> *"it is context sensitive, aware of its temporal, spatial, physical, and virtual environment, operating at a specific time (temporal) and place (spatial), and in response to particular/predetermined real-time triggers (physical) or data input (virtual)."*

AR systems are generally categorised by the following three assets: (1) they combine physical and virtual items in a physical situation; (2) they align physical and virtual items with each other; (3) they function interactively and in real time (Azuma, 1997). These benefits facilitate instructors to overlay virtual graphics over physical items, permitting operators to interact with digital content via physical manipulation. This enables more effective demonstrations of spatial and temporal notions (Wei et al., 2015), as well as the contextual associations between physical and virtual objects.

This contrasts with how VR works, as VR does not need to operate at a 'specific time'. For example, in VWs users can join and rejoin VW and it is still there (persistent) even if their avatar is not. Again, the 'spatial' feature of AR differs from VR, in that VR can bring any 'world' or location to the user, whilst for AR the user must be in that place or world. VR is a structure in which users are totally immersed within an artificial environment and cannot see the outside world. By comparison, AR allows users to perceive virtual items that are overlaid onto or united with the physical world. Therefore, AR complements reality, rather than substituting it entirely (Azuma, 1997).

AR enables the user to be able to move around comparatively liberally in the physical world. These benefits facilitate educators to overlay virtual graphics over physical objects, allowing users to interact with digital content via physical manipulation. Unlike VR, AR does not seek to replace reality, but rather digitally add to it; by combining physical and virtual objects, in real time. Virtual content is combined with the real world in AR (Azuma, 1997). This varies from the concept of VR where the user is immersed inside an artificial environment. In this sense, *“AR supplements reality, rather than completely replacing it”* (Azuma, 1997). It allows a person to see the real, physical world, but it is overlaid with a layer of digital content in real time (Brigham, 2017).

There has been a lack of agreement about one specific aspect of AR, and that is concerning a user’s place in relation to the experience. For example, Samsudin et al. (2014) stated that users of VR are inside that simulated world, rather than peering in from the outside, as with AR. However, this thesis argues that the most significant difference between VR and AR is that users of AR are in the physical environment to which the AR relates, whereas users of VR can be located anywhere.

Bridging virtual and real worlds, AR creates a reality that is enhanced and augmented (Klopfer and Squire, 2008); an example of proven success with such technology is Pokémon GO. The harmony between virtual objects and physical environments allows learners to visualise spatial relationships and abstract concepts, and experience phenomena that are not possible in real life (Arvanitis et al., 2009, Furió et al., 2013). The user is present in the physical environment and views it through a device which enables interaction between that picture of the world and virtual objects.

Three chief affordances of AR: physical world annotation, contextual visualisation and vision-haptic visualisation, were identified by Bacca et al. (2014). A further described affordance of AR is the capacity to connect to outside resources (i.e. the Internet) and extra software on the devices to solve a problem with more success (Klopfer and Squire, 2008). This aligns well with the Connectivist learning theory, which will be discussed later within this chapter (Section 1.8). Additionally, students may control the technologies provided by handhelds in unexpected, yet more efficient ways, e.g. using the video recording feature to make video field notes instead of traditional notetaking (Rosenbaum et al., 2007).

Further affordances include that fact that AR can be used as a creative tool, as investigated by Hobbs and Holley (2016), who aimed to harness the process of creating AR to provide a context for a range of HE skills. Finally, another affordance of AR is the ability to present to a collection of learners numerous incomplete, yet congruent, viewpoints on a problem positioned within a physical space (Squire, 2010, Morrison et al., 2014). This affordance facilitates instructors to integrate cooperative pedagogical methods, such as differentiated role play and the jigsaw technique, which lend themselves to Inquiry Based Learning activities necessitating argumentation.

There are a variety of devices and technologies through which AR can be accessed, but most users experience AR as an application on their smartphone to display media combined with the image of the real world on their screen (Holley and Hobbs, 2020). Typically, the media is triggered by scanning objects or images that match a previously captured image of the scene. Other triggers include dedicated logos, QR codes or the VuMark system (Megali et al., 2006), that trigger the access to and presentation of relevant information. If markers are not accessible, more challenging image recognition (including positional tracking of the user) is needed. One such free application that can be utilised via ‘bring your own device’ (Matt, 2018), is HP (Hewlett Packard) Reveal, which is being widely used in schools and HE Institutions. The Reveal app uses the camera on the mobile device to identify a prompt image. When triggered an 'Aura' (i.e., the pre-recorded media) can be observed on the screen of a portable device.

Like VR, AR has been found (Bacca et al., 2014) to increase student motivation and engagement with the learning process. Mobile AR devices can also enhance the achievement rate of physical interaction-based educational activities and tasks that involve memory – associated learning and support cognitive processing through imagery, and improving information retention (Dunleavy et al., 2009, Rogers, 2011, Sommerauer and Müller, 2014). Improved time-management, verbal communication and enhanced critical reflection skills were also reported by Hobbs and Holley (2016).

Much of the research undertaken has concentrated on identifying the benefits of AR in STEM subjects (Carmigniani et al., 2011, Liarokapis et al., 2017, Enyedy et al., 2012). Other AR studies have been conducted in the fields of Humanities and Arts; e.g.: language learning (Liu and Chu, 2010), visual art (Chang et al., 2016), and culture and multiculturalism (Furió et al., 2013). AR has also been used in Social Sciences, Business and Law, Engineering, Manufacturing and Construction. AR might also be beneficial in health education teaching and nurse training, for instance using AR as part of data analysis and visualisation could help illustrate and interpret results. For example (Hobbs and Holley, 2019), blood circulation data can be projected onto a patient’s body to show blood flow issues from diseases such as diabetes.

Most research performed evaluating AR use in HE has been positive. Several research projects have noted that the interaction between the student and the subject matter improves students' cognitive and learning aptitudes, such as comprehension (Dalgarno and Lee, 2010). The other advantages are the interaction amongst students, and the interaction between the student and the learning tools. These both enable students to detect solutions to problems in a scenario through collaboration. It has been identified that the guiding functions and the sensory experience of these tools can advance students' learning contentment and permit them to complete the learning (Dunleavy et al., 2009, Dalgarno and Lee, 2010). Many investigators (Chang et al., 2010; Johnson-Glenberg et al., 2009) have mentioned the robust prospect of these technologies to engage students of all varieties in immersive practices that improve their instruction.

Attention has surrounded how AR delivers its advantages. For example, Wu et al. (2013) found that AR not only bridges virtual and real worlds but also creates an enhanced reality through a creative process. They argue that the educational values of AR are not solely based on the use of technologies but are closely related to how AR is designed, implemented, and integrated into formal and informal learning settings. One line of argument in this thesis is that it is not the technology itself (in this case VR not AR) that determines learning outcomes but the way in which that technology is woven into the curriculum, including (though not exclusively) how it is related to the learning objectives, the ways in which it can address gaps in student learning, and how it connects to learning assessments. These were all pivotal factors borne in mind when the VR simulation evaluated in the research reported in this thesis was designed and implemented.

A study by Hobbs and Holley (2016) aimed to harness the process of creating AR to provide a context for a range of HE skills within a Collaborative Learning framework. They used the framework to encourage STEM students to improve their 'softer skills' (e.g., communication and problem-solving skills) which are highly desired by employers. The STEM students were required to complete a Personal Development Planning unit of work during their course. The PDP unit was developed by the QAA (2013) to ensure that students are proactive and independent in their studies. Many courses also include lectures about basic academic skills (e.g., presentations, group-work, referencing, report writing, time-management and reflection or self-evaluation). STEM student attendance for these modules was poor as was achievement levels. Hobbs and Holley (2016) felt that a more creative approach to the unit was necessary.

Previous studies (Brown et al., 2008) have shown the value of using fundamentally engaging technology to facilitate group work and stimulate broader skills acquisition. The researchers found an improvement in submission rate between the study cohort and the previous PDP results. Questionnaires demonstrated that the technology was relevant, easy to use and made the course more interesting. Over half of the students felt that the module had improved their soft skills, with approximately one quarter noting improved communication skills, presentation skills, organisational skills and confidence in referencing and writing. This study highlights the creative and student led element of AR, which aligns with the notion of Heutagogy (which will be discussed in Section 1.8).

Despite the various studies that have evaluated AR use in HE, few studies have investigated AR in any systematic way and few large-scale studies have been undertaken. Indeed, Bacca et al. (2014) did not find any studies that used research samples larger than 200 participants. A conceivable reason for this is that larger research samples would need additional devices for a ratio of one-to-one.

Despite AR’s great educational potential, pedagogical constraints of AR have also been reported, including: maintaining overlaid information; paying too much consideration to virtual material; the deliberation of AR as an invasive technology (Bacca et al., 2014); the time-consuming nature of AR tasks (Dunleavy et al., 2009); and covering a certain amount of content within a given time (Kerawalla et al., 2006). Moreover, a successful AR application is reliant upon a capable teacher to facilitate crucial points of the experience (Rosenbaum et al., 2007). The technical limitations of AR include: the challenge of assimilating and handling the AR practice from the educators’ viewpoints (Dunleavy et al., 2009) and Global Positioning System (GPS) error (Rosenbaum et al., 2007). GPS technology is evolving rapidly, at present; however, it concurrently supports and restricts AR implementation.

1.4 The Theory-Practice Gap Identified

Having provided a background to VW/VR and AR, the narrative now progresses to identification of the theory-practice gap that triggered this research. Diabetes education, and the identification and treatment of hypoglycaemia, was chosen as a focus for the VR deteriorating patient simulation evaluated in this thesis. Before moving on to briefly defines some key terms, a short background to diabetes and hypoglycaemia will be provided.

The rationale for choosing this focus derives from a number of reasons including: the statistic that 25% of patients that nurses come across in practice have diabetes as co-morbidity (Diabetes UK Guideline). Additionally, the incidence of diabetes has increased by 60% over the last decade according to Diabetes UK (2019). Globally, the number of people with diabetes mellitus has quadrupled in the past three decades, positioning it as the ninth major cause of death. Approximately 1 in 11 adults worldwide now have diabetes, 90% of whom have Type 2 diabetes (Zeng et al., 2018). The increase in the number of people with diabetes means that there is a greater occurrence of people with diabetes-specific complications, such as kidney failure and peripheral arterial disease (Harding et al., 2019). Furthermore, the epidemiology of other disorders correlated with diabetes, including infections and cardiovascular disease, is also proliferating, with direct effects on quality of life, stress on health services and economic costs (Ahmed and Khan, 2019, Einarson et al., 2018).

Diabetes and its treatment are complex, and for healthcare professionals, inadequate knowledge about diabetes treatments can result in serious hypoglycaemia or hyperglycaemia for the patient (Watts, 2018). Moreover, hypoglycaemia is known to be a complex subject for undergraduates to learn (Chan and Zang, 2007). The most common areas of knowledge deficit include the need to treat blood glucose levels of < 4 mmol/L promptly, to treat hypoglycaemia (all episodes of an abnormally low plasma glucose concentration that exposes the individual to potential harm Ahmed and Khan, 2019). with appropriate carbohydrates and to re-check blood glucose levels after 10-15 minutes (Stewart et al., 2016). These specific areas have been confirmed by the Royal Bournemouth and Christchurch Hospital Diabetic Nurse Specialists, who regularly identify multiple cases of mismanaged hypoglycaemia. Simulations are increasingly being offered as part of the educational experience and valued for their more authentic approaches in preparing for live clinical experience (Bayram and Caliskan, 2019). This thesis argues that VR-based simulation can improve learning about diabetic concepts.

1.5 Previous VR Nurse Education Research

It is felt necessary to scope the previous VR nurse education research prior to setting out the justification for this new study. There is a paucity of quality published literature on the application and/or integration of VR into nursing education (Fealy et al., 2019). As an emerging technology, the literature around VR as an educational intervention is relatively limited. Scoping searches revealed papers relating to its use as a clinical intervention (Juras et al., 2019), in paramedical training (Vaughan et al., 2019) and within nursing and allied healthcare education (Ulrich et al., 2014) , with much of the education and training research focused on high risk, invasive skills such as endoscopy and surgery (Rourke, 2020).

Furthermore, research into healthcare education shows inconsistent evidence regarding both immersive and non-immersive VR simulation use. For example, one of the first meta-analysis (Cook, 2013), indicated that non-immersive VR simulation contributes insignificant differences in knowledge outcomes in comparison to normative instruction (typically lectures).

However, certain areas of VR use for nurse education have led to more synergy in their findings. For example, three articles investigated the relationship between VR simulation and the improvement in awareness of disease, symptom identification and skill competence and were in agreement (Dyer et al., 2018; Kidd et al., 2012; Sweigart et al., 2014). Sweigart and Hodson-Carlton (2013) found significant improvement among student nurses using a psychiatric assessment tool ($p = 0.001$) after they experienced virtual dialogues with patient avatars in the virtual clinic constructed in the world of SL. Students with no virtual interview experience had a lower average score on questions assessed during an interview compared to those who had interviewed a patient avatar in a virtual environment. The data supported the efficacy of VR simulation on the transference of skills to practice for students. This is in line with Kidd et al.'s (2012) study that adopted VR simulation in mental status assessment and communication with patients. The 126 mental health nursing students indicated the benefits of VR simulation as allowing students to learn communication and health assessment skills in a stress-free environment where mistakes would not lead to catastrophic outcomes.

Dyer et al. (2018) also found health profession students demonstrating an increased understanding for older adults who have Alzheimer's disease.

Similarly, a meta-analysis which examined the effect of desktop (non-immersive) VR simulations compared to normative methods of instruction, showed a clear positive net overall effect of learning gains with VR simulations (Consorti et al., 2012). Furthermore, while Davis-Reyes et al. (2008) demonstrated higher cognitive gain in the VR group, this group obtained a lower score than the simulation group in the post-test measurement, showing that although they showed a greater increase in knowledge, their overall test scores were inferior to the simulation group. The opposite effect was seen in a similar study (İsmailoğlu and Zaybak, 2018).

Through a literature review of six articles, Wan and Lan (2019) and Lam (2019) found that improvement in learners' knowledge and attitude was found to be superior in VR group students. However, the technology's improvement in learners' empathy remained disputable among the selected articles. They advocated large-scale RCTs to evaluate VR efficacy for nurse education. Similarly, Kunst et al. (2018) found that inconsistency in the methods used to evaluate VR simulation activities creates challenges in providing definitive answers about the benefits for undergraduate nurse education. Hirt and Beer (2020), having reviewed the literature (six articles) surrounding dementia care and VR learning, concluded that rigorously conducted studies with robust designs are necessary to generate knowledge, and what might be an effective learning strategy in Nurse education. Overall, of the studies that have evaluated the efficacy of VR to support learning in the area of health have been inconclusive in their findings.

1.6 How has the educational use of VR been evaluated?

Before detailing the method of evaluation for this study it is appropriate to examine prior evaluation of the educational use of VR and whether outcomes have been mainly positive, negative or inconsistent overall. Starting with studies that demonstrated positive results, Robert's (2018) study compared the use of a VW to traditional lecture-based teaching via a PowerPoint presentation and demonstrated improved learning outcome scores and improved retention scores after a three-week time delay in the Experimental Group.

In another study, producing positive results, military scholars were trained with either the normative teaching approaches or with a VR-based teaching method (Webster, 2016). They established that the VR group had a higher improvement of 14%. Similarly, Allcoat and von Muhlenen (2018) found that their 99 contributors in both the VR and textbook-style circumstances displayed more improved knowledge than contributors in the video condition. Additional analysis of the learning statistics revealed that participants in the VR condition were superior at 'remembering' than those in the video and old-style conditions, and participants in both VR and old-style conditions were superior at 'understanding' than those in the video condition. These results hint that VR can support not only the lower order of thinking Bloom et al. 1956 cited by Su et al. (2004), but also some higher order thinking including 'understanding' what has been learnt. Indeed, Yildirim et al. (2019) found that VR was effective for long-term knowledge retention, whilst Smith and Klumper (2018) proposed that VR significantly affected short-term retention but did not have an effect of long-term retention. In a similar study conducted by Ogbuanya and Onele (2018) it was established that desktop (non-immersive) VR positively affected students' (n=149) academic achievement and learning interest in electronics technology. Overall, from these studies it is evident that the outcomes have been inconsistent.

Another study that pointed to negative outcomes was that of Parong and Mayer (2018). They found that whilst engagement and motivation were better in their VR group (n=28) compared to a normative instruction group (n=27), learning outcome did not improve. The outcomes indicated that learners who watched the PowerPoint slideshow achieved significantly better on the post-test than the VR group, but described inferior motivation, interest, and engagement scores. This study was an RCT; however, the numbers of students in each group was small. Madden et al. (2019), also found that whilst students (n=172) enjoyed learning about moon phases using VR technology, post-test scores were not significantly better than those who learnt from traditional methods. It was not clear how participants were assigned to groups in this study. Overall, the research discussed up to this point in the thesis is inconclusive in relation to the impact of VR learning.

1.7 Scalability and Sustainability of VR technologies

It has been established that non-immersive computer-based VR simulations might be an accessible and affordable strategy, for delivering flexible and broad-ranging scenarios that focus on cognitive and manual skills in nursing (Dubovi et al., 2017). However, despite the benefits associated with integrating VR technology in education, there are also barriers currently preventing its scalability. These barriers are chiefly related to cost, educators and students learning new IT skills, time to learn how to use hardware and software, additional course preparation time, possible health effects, and facing hesitancy to incorporate VR into the curriculum (Drljević et al., 2017, Moglia et al., 2016).

As with any other new technology, VR could be faced with apprehension when it is applied to new disciplines. Several studies explored how technology adoption and diffusion by faculty members depends on their perceptions, beliefs, and attitudes towards the effect of utilising VR

(Fernandez, 2017). El Hussein et al. (2019) grouped these concerns into the categories of usefulness and acceptance. Additional studies investigated the technical and cultural challenges that arise when using VR when used in education, including cost, usability of software in association with various devices, technology anxiety, and learners' attitude/willingness to incorporate VR in their learning (Huang et al., 2016). Although desktop-powered VR devices have shown promise in recent years, educators have been wary of moving into the mobile space due to the limitations of these devices (Birt et al., 2018).

Indeed, the Horizon Report (2019) reported on the fact that the widespread adoption of VR and AR has been slower than first predicted. However, there are several universities and HE disciplines (e.g., health, medicine and business) who are already delivering authentic learning experiences using such technological approaches to deliver at scale. Noteworthy here is the Australian Catholic University who have worked collaboratively with its School of Behavioural and Health Sciences in the use of HoloLens to explore applications to enhance teaching practice (Horizon 2019).

Universities are increasingly challenged to better equip students with the skills and capabilities that will enable them to succeed in both the current and disrupted future workplaces (Horizon, 2019). Despite previous forecasts for MR's scale in HE, a closer review of trends or approaches that support the creation of increasingly augmented, mixed, hyper, blended, or virtual environments makes apparent that it is not just the technology that needs to be engaged with but also the educational (or learning) outcomes that it is seeking to achieve. If the use of a digital technology is going to be adopted, then it is important that the learning outcome has been defined and the range of approaches that might best enable it to be achieved have been considered.

A further consideration for scalability and sustainability of VR technologies in nurse education is the expense and quality of the technology required. The general simulation pedagogy consensus (Cook, 2013, Zendejas, 2013) is that fidelity is not the most important aspect of simulation pedagogy but rather the design and human affordances. In this thesis, design aspects and human affordances are explored in order to shed light on future scalability and sustainability implications of using VR to support learning in HE. This aspect will be revisited during the discussion and conclusion of this thesis.

1.8 Background to Normative Undergraduate Nurse Education

In the research reported in this thesis VR education was compared to 'normative' education via an RCT. Therefore, a short background to traditional HE education will now follow, including a brief discussion of the learning theories and approaches that will be discussed in Chapter Two as part of the systematic literature review.

There are numerous learning theories and approaches that underpin pedagogy in HE (Hayes et al., 2019), including: transformative learning theory (Smith, 2012), Constructionism (Girvan et al., 2013), Contextual Learning Theory (Chen, 2016), Team-Based Learning (Branney and Priego-Hernandez, 2018), Problem-Based Learning (Parson and Bignell 2017), Cognitivist Learning Theory (Achuthan et al., 2017), blended learning (Stockwell et al., 2015), and the flipped classroom approach (Hack, 2016). Constructivism, Experiential Learning, Situated Learning, Connectivism and Bloom's taxonomy seem most relevant for VR simulation-based nurse education. Hence, a brief overview of these five theories/approaches is now provided.

In Vygotskian social constructivist learning, personal interpretation, decision-making and community cooperation foster long-term understanding and transference of learned concepts. Knowledge is socially constructed; learning is social in nature as in a community of practice; and the learner progresses from novice to expert under the guidance of an expert community of practice members (Dass et al., 2011; Mahon et al., 2010). In short, the construction of knowledge requires learners to be actively involved in the process of learning (Martin et al., 2010). Constructivism does not prescribe planned strategies of instruction but advocates the creation of learning environments that provide opportunities to engage learners in meaning making. Collaboration thus occurs when learners effectively communicate their understanding, engage with the views of others and reflect upon their encounters but does not assume all students are similarly equipped to do so (Proctor, 2019). Some forms of VR, notably VW, lend themselves to collaborative learning (Falconer et al., 2018).

Connectivist Learning Theory is also known as networked learning. According to Siemens and Conole (2011), Connectivism shifts the power in education away from individuals, such as learners and instructors, and onto a collective group. Connectivists assume power in learning can be distributed between three different locations: the instructors, the learners, or the network that forms among all participants. Connectivism closely aligns to learning that harnesses technology that takes the learners beyond the walls of the classroom, e.g., handheld devices and VR. Connectivism is positioned as a new philosophy of education for the digital age, making Vygotsky's concept of zone of proximal development (ZPD) more flexible and stretching it to include learning that lies outside the learner, in social networks and technological tools (Mattar, 2018). It is possible to position it as the development of constructivism in response to the current scenario of the intense use of technology in education, functioning though as a philosophy of education. Mattar (2018) proposes that Connectivism or distributed learning should be considered an updated version of Constructivism, understood as a general philosophy of education for the digital age. Whilst, few nurse education articles (Ross and Cross, 2019) cite Connectivism thus far, it is anticipated that technology-based learning will be increasingly underpinned by this 21st Century approach.

Experiential learning theory provides students with the opportunity to experience a particular aspect of their learning, to be able to observe and reflect on that experience and to form abstract concepts based on their reflections, then apply such knowledge to test new concepts Kolb cited in Baker et al. (2012). Experiential learning theory often underpins simulation-based learning in nurse education. Kolb's experiential learning cycles are viewed as a useful framework within simulation-based learning in health care (Keskitalo and Ruokamo, 2015). In simulation-based education, simulations are viewed as concrete experiences that are debriefed afterwards. This enhances the experiential and reflective characteristics of meaningful learning to be realised in terms of gaining new experiences within the simulation and discussing them afterward (Birt et al., 2018).

Situated learning is a teaching strategy that places the learner in a simulated but realistic environment where they assume a specific role and set of tasks or challenges designed to achieve pre-specified learning objectives. Through situated learning, connections between complex real-world situations and classroom experiences can be developed (Falconer, 2013). One category of situated learning used in pharmacy education is online case scenarios that use virtual patients. Virtual patients may be defined as computer programs that simulate lifelike clinical scenarios in which the learner becomes the healthcare professional making therapeutic decisions (Barnett et al., 2016). Simulations and VR provide the basis for one form of situated learning by modelling specific aspects of real-world complex systems (Yusoff et al. 2011). Users can experiment with the system either by manipulating parameters or participating inside the system and observing the outcomes of their manipulation and participation.

Besides learning theories, there are approaches and models that also guide pedagogic practice. Bloom's Taxonomy model has been widely used in educational settings (Bloom et al., 1956, cited in Martin et al., 2010). According to cognitive complexity level, Bloom classifies human thinking into six levels: knowledge, comprehension, application, analysis, synthesis, and evaluation. Listening to a lecture in class involves mostly recording and recalling information, which are lower levels of Bloom's cognitive hierarchy of learning. In contrast, solving problems in real time during class forces students to synthesise and apply knowledge as they process it (Stockwell et al., 2015). Recently, it has been closely linked with technology integration (Chen, 2016). Huang et al. (2010) claim that "VR is especially helpful when it comes to address issues that require imagination creativity and high problem-solving ability".

Constructivist pedagogy aligns itself closely with Bloom's Taxonomy as it encourages higher order thinking skills such as evaluating and analysing instead of accumulating or memorising (Kaya, 2015, cited in Chen, 2016). Martin et al. (2010) stated that the activities of students during the process of virtual task implementation can be assessed in terms of their success in meeting the objectives, and the alignment between objective, activity and assessment should make VW tasks more integrated, transparent and coherent to both learner and educator.

In addition to these learning theories and approaches to learning, there are epistemological terminologies to consider. Although many methods of instruction exist, some focus more on pedagogical principles (teacher-focused education) and others more on andragogical principles (learner-focused education), there is usually some of each in all methodology (Crosslin, 2016). Knowles (1973, cited in Holmes and Abington-Cooper (2000:1) contrasted the “concept of andragogy”, meaning “the art and science of helping adults learn”, with pedagogy, “the art and science of helping children learn”. In Knowles' view, the teacher is a facilitator who aids adults to become self-directed learners. Many theorists believe the andragogy-pedagogy classification is not perfect, but they cannot agree on a viable alternative either. Moreover, some academics have shunned the word “andragogy” because it is related to the male hormones “androgens”, and therefore goes against feminist views.

Heutagogy is a newer epistemology that combines pedagogy with andragogy to form a modern learning design. Heutagogy can be described as looking “to the future in which knowing how to learn will be a fundamental skill given the pace of innovation and the changing structure of communities and workplaces” (Hase and Kenyon, 2007:2). Concepts that are connected to Heutagogy include self-directed learning, double-loop learning, non-linear learning processes, and learning how to learn. The main idea behind Heutagogy is that learners are not taught what to learn, but how to become a learner in relation to the ongoing learning of a topic or skill set. Most experienced course designers will recognise elements of all three methodologies in almost all classrooms and online courses, though pedagogy prevails.

The normative approach towards teaching hypoglycaemia at Bournemouth University (BU) can be considered as being of a constructivist nature and via a traditional PowerPoint led lecture (with additional reading materials available for students on the Virtual Learning Environment (VLE) – Brightspace™ – an element of Connectivism), with paper-based case studies included to encourage student participation, in an attempt to activate higher levels of Bloom’s taxonomy in the learners. The normative approach for the diabetes session evaluated in this thesis has components of both pedagogy and andragogy. Heutagogy certainly appears to align very well with both AR and VR, though did not underpin the VR deteriorating simulation evaluated in this instance. In future iterations of the software there are opportunities for elements of Heutagogy to be enabled, e.g., by providing more choice and less rigid pathways that the student is expected to pass through to complete the experience.

1.9 Proposed VR Simulation Nurse Education

Nurse Lecturers at BU had requested assistance in innovating an area of the Long-Term Conditions Unit, in which education concerning the treatment of diabetes is located. Examination results and educators' discussions with 2nd Year Nursing students had highlighted the difficulties facing these students in terms of understanding and retaining information about hypoglycaemia (low blood sugar). Furthermore, a Diabetic Nurse Specialist currently working at the local hospital (The Royal Bournemouth County Hospital- RBCH) had liaised with the Nurse Lecturers and highlighted the fact that junior nurses lacked confidence, competence and knowledge surrounding the identification and management of hypoglycaemia. Academics delivering the Long-Term Conditions Unit teach up to 200 students at one time and are timetabled to run the sessions in traditional lectures theatres, hence a scalable and sustainable innovation was required.

VR was identified as a possible means of facilitating learning within the Long-Term Conditions Unit. VR lends itself to Nurse Education because it has been demonstrated to enable repetition of training exercises until competence is achieved. Moreover, VR affords the ability to provide a failsafe and accessible learning environment that may increase patient safety, through repeated exposure to educational content such as clinical skills and critical events as novice practitioners (Butt et al., 2018, Chang, 2016, Cobbett and Snelgrove-Clarke, 2016). VR simulations may provide a more realistic experiential learning environment in the highly flexible and programmable virtual environment to healthcare learners than the conventional simulation (Dalgarno and Lee, 2010).

Virtual reality also has the potential to solve logistical problems, e.g. to simulate learning environments such as home care, operating rooms, or emergency rooms, as well as to bring environments to the learners in a simulative manner. Dangerous care situations or potentially escalating situations may also be addressed by VR simulation. For these numerous reasons, VR was selected, with the aim of spanning the theory-practice gap, engaging and motivating students, and improving their knowledge of the identification and treatment of patients suffering from hypoglycaemia.

To ensure the suitability of the VR simulation, curriculum learning objectives/assessments were reviewed, and research was drawn upon both from Diabetes UK guidelines for the identification and treatment of hypoglycaemia (Diabetes UK 2019), but also in terms of how to design an effective VR simulation. Research shows that a good simulation design integrates mastery of required knowledge, psychomotor skills, clinical reasoning, and reflective thinking skills within authentic scenarios that are either likely to be commonly encountered and/or have significant impact with learning outcomes that can be measured against professional standards for practice (Sapiano et al., 2018). Clear learning objectives when made apparent at the commencement are effective in guiding all aspects of simulation design and in focusing student learning (Kunst et al., 2018). This was considered when communicating to students, via their VLE, prior to the VR session.

1.10 Limitations of Previous Evaluations of VR interventions

Most of the previous VR research has focused on training (e.g., aviation) or therapeutic purposes (e.g., medical) rather than content delivery (Lee et al., 2019). Furthermore, research investigating whether the use of VR technology improves student learning outcomes has been inconclusive. Hence, there is a need for further research to consider the significance of the use of VR in terms of adding value to learning, as research has frequently relied on self-reporting and user experience (Radford et al., 2011, Inman et al., 2011, Radianti et al., 2020). This thesis argues that VR provides unique settings for learning, but the problem lies in how the affordances of VR can be combined in ways to sustain quantifiable advances in learning outcomes.

According to Alfalah (2018) it is crucial to examine the VR system usability by learners, instructors, and by curriculum developers. Little has been revealed about how contextual mechanisms support and interact during such VR learning. Moreover, according to Wimpenny et al. (2012), comparatively limited studies have grasped the numerous individual viewpoints and experiences about student learning. The ways in which such viewpoints affect the instructors' pedagogical design choices and how students make sense of and react to learning in VR have not been evaluated fully (Savin-Baden and Falconer, 2016). Hence, more direction is required on how to make crucial pedagogical use of VR learning. This advice was acted upon in the research reported in this thesis and was used to drive the methodology.

There have been few empirical large-scale studies conducted to examine the efficacy of VR for nurse education, and wider HE, and to drive and direct the sustainability and scalability of such VR technologies (Yung and Khoo-Lattimore, 2019). Moreover, there is a paucity of research that specifically details how VR simulations can be designed to best improve student learning outcomes (Gill et al., 2019, Holder and Bethea-Hampton, 2019, Radianti et al., 2020).

1.11 The Purpose of this Thesis

Bearing in mind the limitations of previous evaluations of VR interventions, the evaluation conducted for this research needed to be more holistic and robust than current evaluative practices. In order to achieve such a thorough evaluation, two literature reviews were conducted. The first review (a systematic review) was conducted to identify successful elements of evaluation that could be utilised in the present research, but also identify areas of weakness in past experimental design. The second review (a CR literature review) was then conducted in order to construct a conceptual model to be tested empirically. This is the reason why two literature reviews were conducted.

The main aim of this thesis was to assess the effectiveness of Critical Realism (CR) paired with Partial Least Squares-Structural Equation Modelling (PLS-SEM) as a method to evaluate the impact of VR simulations on undergraduate nurse education. The aim has been achieved through two objectives, which were:

1. To determine the effect of pairing CR with PLS-SEM as an evaluative method

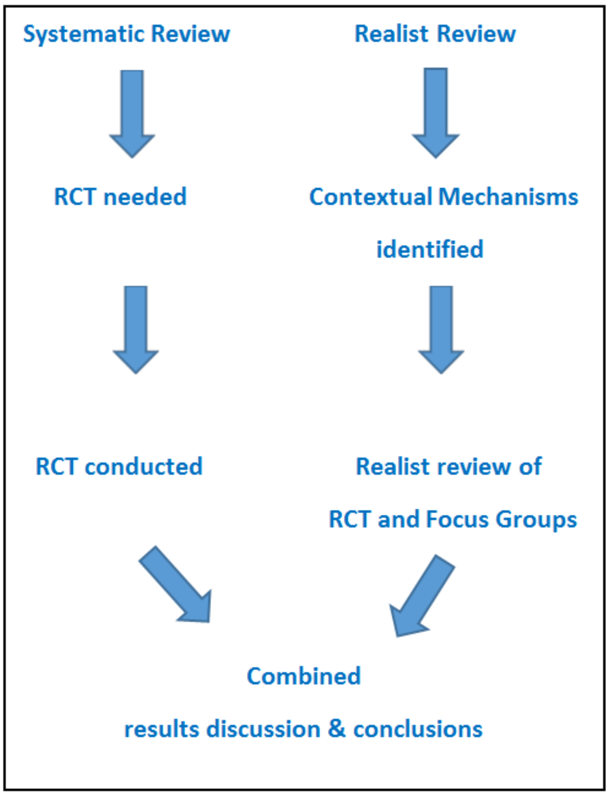
2. To determine how using this novel evaluative method can inform our understanding of the impact and future use of VR simulations for undergraduate nurse education

This study sought to obtain data to address these objectives. If we understand the practical advantages and challenges surrounding VR use in a large lecture theatre setting, from the point of view of students, academics and learning technologists (LTs), we might better understand how VR activities can be interwoven with current student nurse education and begin to understand how such practices can become both scalable and sustainable in the future.

1.12 Research Design and Methods

This study used first a systematic review to scan the current literature (Chapter Two), followed by a CR approach (Chapter Three) to investigate whether VR learning can complement traditional (normative) methods and in what circumstances it can be used. Qualitative and quantitative research designs were adopted to provide a holistic evaluation of VR use in nurse education. The research data in this thesis was drawn from a student survey and focus group discussions (FGDs) with students and staff at BU. Figure 1-1 illustrates the method sequence.

Figure 1-1 The Method Sequence used in this Thesis



1.13 The significance of the current study

Accepted findings and theories of other researchers were built upon and then extended beyond previous knowledge by incorporating original elements. In response to gaps identified through a systematic literature review, knowledge was fed forward to design and conduct a robust RCT. A VR learning conceptual model was proposed (containing some original pathways) to test whether significant "causal" relationships existed between identified latent variables. Understanding the link between the variables that interact when VR learning takes place will help HE academics, course designers and LTs who are planning to use VR to support teaching and learning within the HE sector, in particular in the area of nurse education, by designing and implementing a new methodological approach.

This study also offers some important insights into pedagogic knowledge relating to the implementation of VR exercises within traditional lectures from students', academics' and LTs' points of view. This research is original methodologically, as it is the first of its kind to combine both an RCT with a CR angle for evaluating a VR intervention in HE. The pairing of PLS-SEM with critical realism is unique to this study (previously only SEM has been combined (Ford et al., 2018). This project provided an important opportunity to advance the understanding of the use of VR learning in HE settings in terms of efficacy, scalability, sustainability and transferability considerations.

1.14 Chapter Summary

To recap, AR is most appropriate for HE teaching when students can be in the learning environment they will be interacting with. AR enables learners to layer digital aids and visualisations onto physical objects/settings. If a learner needs to bring a learning or environment into their traditional lecture theatre then VR would be more suitable, because it substitutes the physical world with a computer created one. The VW lends itself to collaborative learning whereby students and academics can enter the VW at agreed times and complete ongoing activities. VR was deemed to be more appropriate for isolated simulation activities that are required to take place in a large-scale lecture theatre setting.

In this chapter the scene was set for the thesis, clarifying the technical terminology that will be used throughout, outlining historical beginnings and developments in VW/VR/AR, and discussing the affordances and barriers to the adoption of such technology use in HE and other settings. The chapter concluded by setting out the rationale for the research and a brief justification for the methods and approaches used to design, implement and evaluate the VR intervention. An integrated systematic review of AR and VW/VR use in HE will now be presented in Chapter 2.

Chapter 2 Systematic Review of the Literature

2.1 Synopsis

Increasing availability of immersive VR, including VWs and AR equipment for use in educational contexts, are providing new opportunities for instructional designers (Ritz and Buss, 2016). Educationalists endeavour to find guidelines that could assist them in developing and evaluating original innovative learning systems, because they need assurance that such practices are superior to traditional methods. According to Quinney et al. (2017), responding to the changing landscape of HE requires the development and implementation of imaginative approaches to constantly inspire, engage and support staff in delivering high quality student-centred learning experiences. This chapter evaluates how different approaches to researching these issues can give us different perspectives and makes recommendations for appropriate methods for different forms of evaluation. It is argued that a multi-methodological approach therefore has merit in researching this area. Additionally, it is reasoned that, although the rigour of study design is steadily improving, stronger evaluation is still required to investigate the efficacy of the VW/VR/AR technologies that are used in HE.

In this chapter the outline of the rationale for undertaking an integrative systematic review is set out, and the review protocol and methodology are described. The chapter proceeds to report the findings of the review, exploring reported VW/VR and AR practices in HE along with underpinning theoretical schema, aligned pedagogies and how the use of such technologies has been evaluated within the articles included in this review. The chapter ends by describing how the findings of the review influenced the empirical part of this thesis.

2.2 Rationale for conducting a systematic review

Studying prior research in a field is vital, as this exposes the existing state of the field and proposes direction to researchers who are looking for appropriate topics to discover (Akcayir and Akcayir, 2017). Petticrew and Roberts (2008) demarcated a systematic literature review as an interpretation of a selection of documents on a topic that optimally comprises summarisation, analysis, appraisal, and fusion of the papers. The benefit of such a systematic review is that it yields a map of the 'bigger picture'. However, the method has been criticised as taking a reductionist angle on study evidence, possibly leading to restricted findings (MacLure, 2007), though the current tendency towards including a vigorous qualitative slant (Higgins and Green, 2011) has gone some way towards addressing this apparent constriction of scope.

Systematic reviews can enable comprehension of a topic, detect common threads through studies, and aid in the growth of theory (Tondeur et al., 2012), hence selection for use in this thesis. Furthermore, it is a fundamental requirement of all doctoral research that the research demonstrates an original contribution to current knowledge. It is only by undertaking a thorough review of the current evidence, preferably through a systematic review, that current gaps in the knowledge base can be identified.

Virtual technologies that immerse the user in a simulated world or enable the overlay of simulated artefacts or environments onto our current view of the world, have become increasingly accessible over the past 10-15 years. The variety of equipment and applications available continues to grow rapidly (Maravilla et al., 2019) and these technologies have now been absorbed into the range of recognised resources available to educators, particularly those wishing to use simulation and collaboration in their education practice. This technology proliferation is creating exciting opportunities in educational contexts. However, HEs, during a time of economic uncertainty, need a reliable evidence base before investing in learning technologies. This, in turn, raises questions for educators, which might include: Which virtual technology would be most appropriate for my subject and my students? How might using this technology affect their learning outcomes? What barriers would I have to overcome to use this kind of technology? Educators wishing to use virtual technologies in their practice therefore need access to research findings that can assist them in evaluating how they might do this effectively, helping them to reconceptualise their teaching, and their students' learning. There is a developing research literature on virtual technologies, which reports upon and evaluates their use in a range of educational settings in HE. This thesis responds to an apparent gap in this literature, i.e. how different methodologies might influence the findings of research in this field.

This research has a strong rationale considering the rapid evolution of VW/VR and AR tools, which has not been accompanied by the systematic evaluation of efficiency and effectiveness for learning. Whilst there is a wealth of qualitative and descriptive studies delineating discrete interventions using VW, more research is vital to equate the different pedagogical methods that can be provided in-world and their influence on pupils' learning (Hack, 2016). Enhanced student engagement and wide applications for VR/VW use in HE has been frequently reported in the literature (Philips et al., 2015), but research has often relied upon self-reporting for learning outcomes (Hew and Cheung, 2013). The definitions, affordances and barriers of VR/VW/AR learning systems have been systematically reviewed, but gaps in the effective evaluation of these technologies still remain (Yilmaz and Batdi, 2016, Akcayir and Akcayir, 2017, Fowler, 2015), particularly evaluation of AR technologies (Martin-Gutierrez et al., 2014, Munnerley et al., 2014).

What is already known about this topic?

- Definitions and descriptions of VR/VW/AR learning systems
- Learning affordances/benefits of VR/VW/AR learning systems
- Barriers to implementation of VR/VW/AR learning systems

2.3 Scoping Review

A scoping review was conducted in order to both direct the search string and gain insights about the topic. As a result of the scoping review, 11 literature reviews were retrieved from the BU library system. Of these reviews, five had reviewed 0-50 articles, three had reviewed 51-100, two had reviewed 101-150 articles, finally one had reviewed 165 papers. Table 2-1 summarises the findings of these previous reviews.

Table 2-1 Scoping Review Main Findings

Short title	**Sample size**	**Focus**	**Main Findings**
Dass et al. (2011)	15	VWs	Students generally receptive to the use of VWs Disadvantages Concerns related to technological requirements and cooperative and communication tools
Inman et al. (2011)	27	VWs (SL)	Student-centred learning
Mikropoulos and Natsis (2011)	53	Educational Virtual Environments	Studies present real-world, authentic tasks that enable context and content dependent knowledge construction Disadvantages Little can yet be concluded regarding the retention of the knowledge acquired in EVEs Recommendations Longitudinal studies
Duncan et al. (2012)	100+	VWs	Aids higher order thinking and interactivity Disadvantages Distractions in the VW; VW is not enough to improve cognitive outcomes
Wang and Burton (2013)	107	VWs SL	The educational implementation of SL is no longer in its infancy; research studies on SL have been widened and deepened.
Steils et al. (2015)	90 FGDs	VWs	What is needed is a mapping of VW theories, pedagogies and practices so that it is possible to delineate their impact. This will enable those using or wanting to develop VW learning to be clear(er) about the purpose, practices and pedagogies involved

Ghanbarzadeh et al. (2016)	165 papers over 10 years	VWs	Collaboration, team work and self learning. <u>Recommendations</u> Longitudinal studies. International studies
Kurilovas (2016)	33	VR/AR/MR	<u>Recommendations</u> In order to fully evaluate the quality of the learning platform/environment, both expert-centred (so-called top-down) and learner-centred (bottom-up) approaches should be applied
Moglia (2016)	38	VR simulators	<u>Disadvantages</u> The cost-effectiveness of VR simulators compared to other training approaches is largely unknown. Overall there is no evidence on the transfer of skills gained using virtual simulators to the operating room. <u>Recommendations</u> RCTs, preferably multicentre, are required to solve this issue, with the goal of facilitating the adoption of VR simulators in curricula for robotic surgery.
Drljevic (2017)	59	AR	<u>Disadvantages</u> Most ARLEs do not provide effective consideration for the techno-pedagogical design needs of the teacher or other facilitator in order to enable them to effectively manage a class through reduced orchestration load
Pellas (2017)	33	VWs	<u>Advantages</u> VWs can influence students' attendance, knowledge transfer, skill acquisition, hands-on digital experience and positive attitudes in laboratory experimental exercises and opportunity to manage learning materials more effectively and efficiently during the teaching process.

From this scoping review, the two most relevant reviews were Ghanbarzadeh and Ghapanchi (2018) and Kurilovas (2016). Ghanbarzadeh and Ghapanchi (2018) analysed 165 VW papers over a 10-year time period (hence its significance), concluding that VWs can play a substantial role in improving interaction among students and educators. The limitation of this work is its scoping, which refers to VWs. Kurilovas (2016) synthesised 33 VW/VR/AR articles (covering all technologies relevant to this thesis); but was hampered by a limited search string and a single database.

Both studies advanced knowledge surrounding affordances of the technology; however, evaluation was not a key theme in Ghanbarzadeh and Ghapanchi's (2018) study. Kurilovas' (2016) research identified two recurrent limitations in previous studies; these were that experiments have been mainly evaluated with small to medium heterogeneous samples, secondly that learning platforms were evaluated by the users only, a drawback also acknowledged by Merchant et al. (2014). This review aims to build upon the work of these two papers and they will be revisited in the discussion section of this chapter.

A slightly earlier systematic literature review was conducted by Mikropoulos and Natsis (2011) and studied over 50 papers, straddling 10 years (1999-2009), regarding the use of VR. One of their interpretations is that very few (n=10) of the studies reviewed had a transparent conceptual model to enlighten curricula design. Where a theoretical model was offered, it was nearly always grounded in constructivism, often implicit, or was a variation of the method (e.g., PBL, experiential learning, or collaborative learning). "All the other reviewed articles do not refer explicitly to a learning theory" (Mikropoulos and Natsis, 2011). There is therefore a space in research concerning theoretical pedagogical models to inform these systems. This paper was written nine years ago and therefore the technologies evaluated might now be outdated; however, the need for clarity surrounding pedagogical underpinnings remains.

Linked to this idea of the importance of pedagogy when evaluating VR use is the detailed examination of VWs by Steils et al. (2015), (though not a systematic review). They reported a large-scale study integrating three doctoral papers which discovered the socio-political impact of T&L in VWs. Findings indicated that “liquid learning” and “liquid curricula” can guide the way technology-enhanced learning should be assimilated into HE. Steils et al. (2015) recommended that future work on VWs needs to take account of the experiences of tutors and students. They warned that change has created curricula with diminutive pedagogical underpinning and a tendency towards technological determinism. Similarly, Selwyn (2010) believes that much attention has been given to promoting the use of educational technologies, though less focus has been on exactly how they have been used and their respective impacts. Steils et al. (2015) called for a charting of VW theories, pedagogies, and practices in order to define their impact. Key points were taken forward from this review to inform the systematic review of the literature.

This scoping review established two things: firstly, that there is enough original research to undertake a systematic review and secondly, to ensure that the systematic review planned has not already been completed. Both criteria were confirmed; hence the systematic review was conducted (as described below).

2.4 The Systematic Review Process

Enhanced student engagement and wide applications for VR/VW use in HE have been frequently reported in the literature (Englund et al., 2017, Philips et al., 2015, Warburton, 2009, Matthew and Butler, 2017). The definitions, learning affordances and barriers, e.g., McFaul and Fitzgerald (2019), of VR/VW/AR learning systems have been reviewed, but gaps in their effective evaluation remain (Yilmaz and Batdi, 2016, Akcayir and Akcayir, 2017, Fowler, 2015, Baxter and Hainey, 2019, Marcel, 2019). This is particularly evident in evaluations of AR (Martin-Gutierrez et al., 2014, Munnerley et al., 2014, Garzon and Acevedo, 2019, *DePape* et al., 2019).

This thesis has also established that there is also a gap in the literature concerning theoretical models to inform these systems (Mikropoulos and Natsis, 2011, Rotkonen et al., 2019). This thesis reflects upon the limitations of the two most significant systematic reviews in the past three years (Ghanbarzadeh and Ghapanchi, 2018, Kurilovas, 2016). Building upon these reviews, the aim of this chapter is to systematically review the literature across the three virtual technologies discussed above, in order to identify evaluation approaches that will be of use to education practitioners seeking robust evidence. Hence the following two questions have been addressed in this systematic review:

What underpinning learning theories and practices have been reported for VW/VR/AR learning systems in HE?

How has the effectiveness of VW/VR/AR learning systems been evaluated?

Within this section a description will be provided relating to the systematic review process used and why specific methods were chosen. This study was conducted as a integrative systematic review of a body of knowledge according to guidelines proposed by Kitchenham (2004), whose review framework is one of the most extensively used and consistent approaches (Ghanbarzadeh and Ghapanchi, 2018, Yadegaridehkordi et al., 2019). Kitchenham (2004) describes a systematic literature review as a methodical approach for identifying, evaluating, and interpreting previous studies conducted in a specific research field.

Systematic reviews are mainly either meta-analyses (reviews of quantitative data) or meta-syntheses (reviews of qualitative data) (Athanasakis, 2019). Integrative systematic reviews include both qualitative and quantitative data (Løyland et al., 2020). Hence, the review approach taken within this thesis is an integrative systematic review.

The stages in research methodology are described, and the processes of paper inclusion and exclusion as well as extracting and analysing data are identified. For inclusion, articles needed to be published within the last 10 years, be peered reviewed and written in English. Original peer-reviewed research papers conducting qualitative, quantitative or mixed designs were included, but not systematic reviews. In this study, the articles from journals, thesis, conferences and workshops published in the English language from January 2009 to August 2019 were considered. There were reasons behind selecting this period in this study. For example, this review is a complement to previous review studies and provides a more in-depth understanding of affective computing in education in recent years. However, there is no effort that specifically focuses upon and reviews studies published from 2009 onwards.

The search excluded studies that were: not written in English (as the author of this thesis would be unable to evaluate them), papers that did not focus upon HE, and papers that exclusively focused on educating students with special needs, as these would add an extra level of complexity depending upon the specific needs that were being catered for. Tables 2-2 and 2-3 summarise the inclusion criteria and excluded terms.

Table 2-2 Inclusion Criteria used for the Systematic Review of the Literature

Criterion Type	**Inclusion Criteria**
Relevance	Literature must relate directly to empirical work focusing on the RQs
Recency	Literature should have been published between 2009-2019. 2009 corresponds to the introduction of Oculus Rift which galvanized the launch of other similar technologies.
Age-range	Literature must relate to the HE age range (which might include colleges e.g. in the US).
Geographical Spread	Literature can be from any country as long as it is published in English.
Research Base	Literature must be based upon empirical research either qualitative or quantitative.
Transparency	The methodology must be explicit e.g. sample size, instruments and analysis
Reliability/validity	Taking into account the type of study the findings must be determined (as far as possible) to be valid and reliable.

Table 2-3 Excluded Terms used for the Systematic Review of the Literature

Justification for Exclusion	**Term/s**
Ambiguity	HE, train* or trainer*
Might locate non-HE articles	teacher
Yielded irrelevant articles	Undergraduate/postgraduate
Tangential	"web-based simulation", "simulated environment", haptics, "3D CAVE*", CPD, "Artificial intelligence" OR "AI", MOOCS, VLEs

The six-stage process for searching for appropriate literature was performed according to Kitchenham's (2004) guidelines and has been represented as a PRISMA (Moher et al., 2009) diagram in order to aid transparency for the reader. In the first stage, 18 digital databases were chosen (as advised by the BU librarian) for the search of keywords, with most articles deriving from Education Source, Academic Source Complete and ERIC. Each source was piloted to ensure that search terms were relevant. It is believed that the choice of these 18 libraries (Table 2-4) was justifiable because together they represent thousands of peer-reviewed journals and conferences. The search strategy was created in consultation with a health sciences librarian with expertise in systematic review searching.

Table 2-4 Database sources- showing number of articles retrieved before duplicates were removed

Database	**No. of Articles**
Education Source	225
Academic Search Complete	135
ERIC	102
ScienceDirect	40
Directory of Open Access Journals	28
Business Source Complete	22
Library, Information Science & Tech Abstracts	21
IEEE Xplore Digital Library	18
SwePub	15
CINAHL Complete	14
Communication Source	12
British Library EThOS	10
JSTOR	6
SocINDEX with Full Text	4
MEDLINE Complete	2
SciTech Connect	1
Emerald Insight	1
Infomit Humanities & Social Sciences Collection	1

In Stage 2, the SPIDER question format was utilised to build the search strings, as it is suitable for searches that combine qualitative and mixed methods articles. Questions based on this format identify the following concepts: (1) **S**ample, (2) **P**henomenon of **I**nterest, (3) **D**esign, (4) **E**valuation, and (5) **R**esearch type (Cooke et al., 2012). During piloting **D**esign retrieved more relevant papers than **E**valuation, hence its inclusion in the search string. Piloted search terms were reiterated following citation mapping and consultation with an expert panel. The advanced search services provided by each of the digital libraries were used to search a combination of title, abstract and keywords of the papers by applying search terms. The first search was performed on the 5th January 2018, and was repeated monthly to check for new publications, until the end of September 2019. Table 2-5 shows the search strings that were used. Boolean techniques were used to safeguard that no relevant literature was missed in the search strategy (Gerrish and Lathlean, 2015). To ensure quality, a second reviewer (Falconer 2017) was enlisted to check a sample of 100 papers against the inclusion and exclusion criteria.

Table 2-5 Search String used for the Systematic Review of the Literature

SPIDER TOOL	Search Terms	List other related words or terms
S	**Search term 1:** student*	**OR** "university student*" **OR** learner* **OR** lecturer* **OR** educator*
P of I	**Search term 2:** "virtual reality"	**OR** "Immersive virtual reality" **OR** "virtual world*" **OR** "immersive virtual world*" **OR** "mixed reality" **OR** "immersive virtual environment*" **OR** "immersive environment*" **OR** "immersive world*" **OR** "virtual reality platform*" **OR** "3D virtual simulation" **OR** "3D virtual learning" **OR** "educational virtual environment*" **OR** "augmented reality" **OR** "augmented reality applications" **OR** "Second Life"
D	**Search term 3:** questionnaire*	**OR** survey* **OR** interview* **OR** "focus group*" **OR** "case stud*" **OR** review **OR** "systematic review" **OR** "literature review" **OR** stud* **OR** "longitudinal stud*"
R	**Search term 4:** quantitative	**OR** "mixed methods" **OR** "RCT" **OR** "trial" **OR** "longitudinal" **OR** "cross sectional" **OR** "action research" **OR** "qualitative"

In Stage 3, all the papers' titles were reviewed one by one and this excluded 151 papers with completely irrelevant topics, according to their title. In Stage 4, the collections of papers were refined by reading the abstracts of the remaining 388 papers and eliminating a further 188 irrelevant papers.

In Stage 5, full texts of the remaining papers were read and excluded 187 based on irrelevant full text. Articles meeting the inclusion criteria were assessed for quality using the CASP tool (CASP, 2013) and with a checklist for prevalence studies from the Norwegian Institute of Public Health (NPH) based on EBMH Notebook (EBMH Notebook, 1998; NPH, 2017). CASP was used for assessing qualitative studies in the selection, with a maximum score of 10. The checklist includes the following criteria: (a) Was there a clear statement of the aims? (b) Is a qualitative methodology appropriate? (c) Was the research design appropriate to address the aims of the research? and (d) Was the recruitment strategy appropriate to the aims of the research? Further, (e) Were the data collected in a way that addressed the research issue? (f) Has the relationship between researcher and participants been adequately considered? (g) Have ethical issues been taken into consideration? (h) Was the data analysis sufficiently rigorous? (i) Is there a clear statement of findings? and (j) How valuable is the research?

Quantitative studies were assessed for quality using consideration of sampling strategy, cohort/sample size, experimental design (e.g., pre-testing and use of a control) and who the research had been funded by (Cohen et al., 2017). No studies were eliminated on these grounds.

In Stage 6, at the end of the search and selection procedure, the total number of selected papers was 50. The reasons for exclusion at the abstract, full text and quality assurance phases are displayed in Figure 2-1.

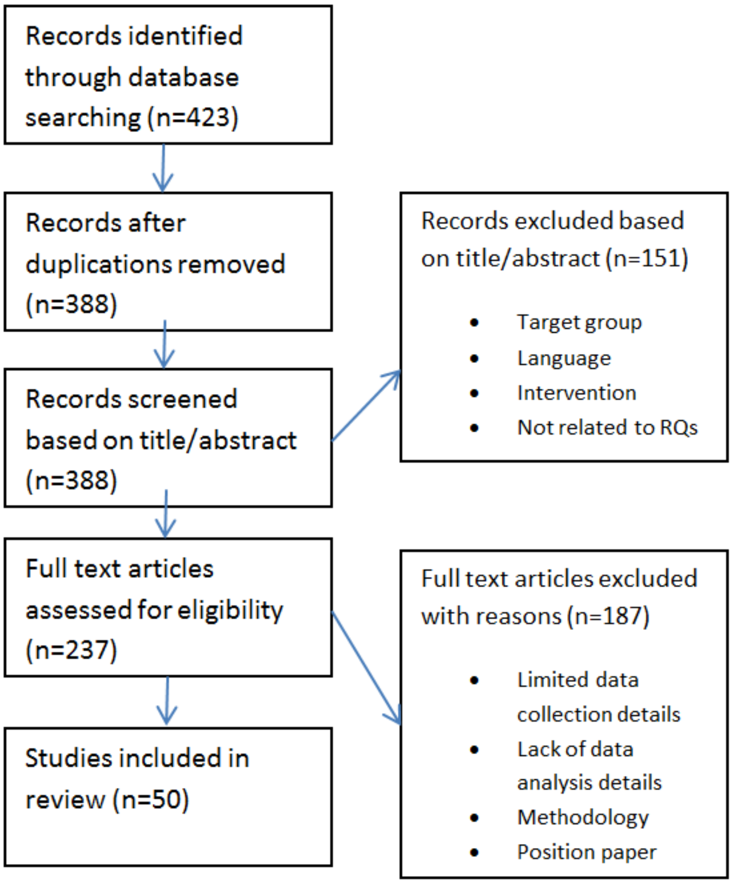


Figure 2-1. Selection flow chart showing reasons for excluding articles during the systematic review of the literature

Three broad types of data were extracted from the studies:

1. VR/VW/AR technology practices and their applications to HE
2. Underpinning learning theories
3. Types of evaluation used

Data analysis was conducted by reading the full text of all 50 papers, identifying themes and synthesising findings related to the research questions (RQs). Analysis was conducted by means of thematic synthesis, a method that is endorsed when the findings have application to practice and policy (Booth, 2016). EPPI Reviewer 4.0 software was used for screening, data extraction, quality assessment, and to conduct the synthesis (Gough et al., 2012). Findings and themes from qualitative papers (n=15) have been incorporated into the findings and discussion section, especially concerning the advantages and challenges of the technologies involved. The efficacy of the use of technologies was evaluated in a range of different ways in the quantitative and mixed methods papers. As a starting point to assess and decide on the strength of approach classifications for strong study designs based upon Ader et al. (2008), Creswell (2018), Cohen et al. (2017) and Palinkas et al. (2015) were used, and these details are reported in the "methods used" part of Table 2-5.

2.5 Principal Findings of the Systematic Review

The 50 included papers are summarised in Table 2-5.

Table 2-5 Principal Findings of the Systematic Review of VW/VR/AR Literature

Paper	**Type of Intervention**	**Data Collection**	**Methods used to ensure rigour**	**Participants**	**Main Finding/s**
Chen (2016)	VW	experiment using SEM	pre-test objective	n=448	Student cognition improved through immersion, ease of use and help-seeking affordances of VWs.
Chodos et al. (2014)	VW	movement analysis	balanced objective	n=2	VW is useful in an emergency medical education context (small case study).
Claman (2015)	VW	quasi-experimental two group post-test	control balanced significant	synchronous method-n=10 asynchronous-n=11	Synchronous learning platform led to significantly higher engagement scores.
Englund et al. (2017)	VW	semi-structured interviews	na	n=12	More than half of the students enjoyed using VWs and saw the advantages of having a setting where communication can be practised in an authentic but 'safe' environment available online. But many students also reported technical difficulties.
Girvan and Savage (2019)	VW	open non-directive interviews, chat logs, constructed artefacts, learners' written reflections and observations	na	n=24	Concluded that VWs are effective environments for constructionist learning.
Gustafsson et al. (2017)	VW	online course evaluation	balanced objective	n=42	Students valued the use of VWs in clinical pharmacy teaching; however, there is a need to make the virtual environment more realistic and easier to use.
Hack (2016)	VW	module failure rate measurement	heterogeneous	n=47	Following the introduction of the VW, the module failure rate was less than 5%, compared with ~11% in the previous two

			balanced objective		cohorts. However, technological challenges to both staff and students were raised.
Jee (2014)	VW	negotiation of meaning model was used as the main construct to analyse students' language use	na	n=34	Data pointed to the utility of a sociocultural approach and led to a focus on patterns of negotiation of meaning as differentially elicited in different task types and on an in-depth analysis of the students' collaborations with their group members.
Mathews et al. (2012)	VW	survey lecturer observations and reflections	balanced objective longitudinal heterogeneous	n=97	Not all students are inclined to take part in technologically delivered education experiences. Technological difficulties and student perceptions of relevance were identified as challenges.
Matthew and Butler (2017)	VW	survey	heterogeneous balanced	n=363	Greater student engagement and improved learning outcomes with VW.
Mayrath et al. (2010)	VW	survey interviews observation 2-year pilot study longitudinal	longitudinal heterogeneous balanced objective	n=150	Student survey and interview responses showed evidence of immersion, timelessness, sustained engagement, effortless concentration, and enjoyment.

Minocha and Reeves (2010)	VW	surveys interviews	pre-test	n=46	VWs are being utilised to foster creativity among students, aid socialisation, facilitate informal learning and enable exploratory and experiential learning rather than traditional instructional ones.
Parson and Bignell (2017)	VW	surveys	heterogeneous	n=19	Educators reported that the experience provided more immersion and engagement than traditional methods.
Pellas (2016)	VW	survey	heterogeneous balanced	n=95	Engagement of students improved, and it was found to be useful for distance learners.
Quintana et al. (2017)	VW	3-year study observation grids and personal log books	heterogeneous balanced longitudinal objective	n=10	Students improved their technology skills and educational aspects about good practices in classes.
Saiya (2017)	VW	surveys 1-year study	heterogeneous longitudinal balanced	n=180	Long-term learning gain of a VW (*Statecraft*) was disputed. Objective and heterogeneous studies were recommended.
Salmon et al. (2010)	VW	semi-structured interviews and from chat log	heterogeneous longitudinal	n=12	Students achieved knowledge construction through following up each other’s questions, keeping the discussion growing, and sharing and exchanging views. Individualised learning was also possible.
Savin-Baden et al. (2010)	VW	semi-structured interviews face-to-face, by telephone and in-world	na	n=20	Learning in groups in a VW has value.

Savin-Baden et al.(2011)	VW	interviews longitudinal	na	n=10	Feedback suggested that the information-driven scenarios did not work as well as avatar-driven.
Savin-Baden (2013)	VW	interviews	na	n=7	Students appreciated the value of a VW as a collaborative environment, but also viewed such practice-based simulations as valuable for individual work.
Sunnqvist et al. (2016)	VW	interviews	na	n=24	A VW promoted students' independent knowledge development, critical thinking, reflection and problem-solving ability.
Wimpenny et al. (2012)	VW	interviews observation longitudinal	balanced longitudinal objective heterogeneous	n=70	Using VWs for learning in particular ways can help students to learn about the cultural values of their discipline and signature pedagogies. However, students might not see the relevance of VWs to their learning.
Yi Fei et al. (2017)	VW	surveys RCT	control objective heterogeneous significant randomised	n=80 randomly assigned to four Experimental Groups	Chatbot and time machine increase the learners' sense of immersion and presence. Best design practices should address how immersion and presence can be integrated into affordances of VWs.
Yu-Chih et al. (2010)	VW	surveys	balanced significant	n=42	Flow experiences in VWs had a significant and positive impact on students' attitudes towards e-learning.

Achuthan et al. (2017)	VR	post lab exercise	pre-test control balanced heterogeneous objective significant	control group-n=45 experimental group-n=145	The VR group had significantly superior learning scores compared to the Control Group.
Barnett et al. (2016)	VR	surveys submitted subjective/objective assessment plan	pre-test control balanced significant objective	control-n=553 experimental-n=581	VR increased student engagement.
Birt et al. (2018)	VR	RCT surveys interviews observation	pre-test control balanced heterogeneous objective longitudinal	n=26	VR increased learner motivation and skills, but pedagogic challenges were also highlighted.

Bower et al. (2017)	VR	qualitative analysis of participant evaluations	na	n=45 Heterogeneous	Insights into supporting and constraining factors for collaborative learning were identified.
Bridge et al. (2014)	VR	RCT	control balance significant heterogeneous	n=48 completed the test	VR group performed better.
De León (2013)	VR	mixed methods	pre-test	n=28	Most participants exhibited some transformation.
Dooley et al. (2014)	VR	qualitative case study	na	n=18	Pre-service teachers were able to learn about a technology integration activity within the context of building English language arts pedagogical content knowledge.
Foronda et al. (2016)	VR	Mixed methods	balanced objective	n=54	Most students suggested that the virtual simulation was a positive experience. However, more research was recommended.
Gamage et al. (2011)	VR	in-depth semi-structured interviews	na	n=22	Perceptions of VR affordances for learning by educators with no experience in using VR are like the perceptions of early adopters and are overall positive.
Girvan et al. (2013)	VR	interviews, chat logs recorded and learners' artefacts and reflections collected	na	n=24	SLURTLES provide learners with a programmable, low-floor, high-ceiling and wide-wall construction tool, which supported their construction of a wide range of complex artefacts as part of a constructionist learning experience.
Hee Lee and Shvetsova (2019)	VR	mixed methods FGDs and survey	control balanced	experimental group n=62	Improved competency with VR.

			objective significant heterogeneous	control group n=51	
Keskitalo (2011)	VR	(gloved) model was used to evaluate students' meaningful learning experiences	na	n=54	Results suggest that VR supported the process characteristics of meaningful learning and its outcomes, although the individual, critical, and interactive characteristics were not fully realised.
Keskitalo and Ruokamo (2015)	VR	case study mixed methods	balanced objective	nine facilitators and 25 students	The analysis of the post-questionnaires suggested simulation-based learning was very meaningful for the students.
Mahon et al.(2010)	VR	surveys	balanced	n=16	Students found the simulation to be a useful learning experience and put them in situations that forced them to think on their feet.
Martin et al. (2010)	VR	measurements of flow, video, text/chat, diaries, and interviews	objective heterogeneous	Teams n=unknown	One of the significant educational benefits of using VR is that it allows individual and cooperative convergence as well as collaborative learning to take place.
Peterson-Ahmed (2018)	VR	exploratory case study using randomisation	control balanced longitudinal randomised	n=8	Coupling VR simulations with traditional teaching methods, allows for increased and individualised remediation for pre-service teachers.

Sattar et al. (2019)	VR	survey	heterogeneous control significant objective	n= 87	Improved learning motivation and learning competency with VR.
Tiffany and Hoglund (2016)	VR	two written assignments qualitative analysis	na	n=15	Students indicated that they increased their own capacity to understand, appreciate, and relate to people different from themselves.
Verkuyl et al. (2020)	VR	mixed methods- FGDs and survey (pre- and post-test)	balanced	n=127	VR leads to positive outcomes in self-confidence and satisfaction.
Quang et al. (2015)	VR and AR	surveys	balanced	n=20	Using mobile based VR+AR would improve construction safety & health effectively.
Castillo et al. (2015)	AR	survey	0	n=59	AR has the potential to be a valuable complementary teaching tool for topics that benefit from contextual learning experience and multipoint visualisation, such as the quadratic equation.
Fengfeng (2015)	AR	control mixed method	pre-test control balanced objective	n=34	AR tended to promote better componential competencies of technological pedagogical knowledge.
Martin-Gonzalez et al. (2016)	AR	survey	objective balanced	n=18	Most users had a positive attitude towards using our AR system for their learning in Euclidean vectors.

Rizov and Rizova (2015)	AR	knowledge and skills test survey	heterogeneous balanced objective significant	n=333	HE educators found that using AR is significantly improving the learning process of students and their teaching process in a pedagogical and technical sense.
Turkan et al. (2017)	AR	control and test groups are deployed, and students' performance is measured using pre- and post-tests	pre-test control balanced objective	Control-n=19 Experimental-n=22	AR can support constructive engagement and retention of information in students.
Yusoff et al. (2011)	AR	regression analysis	heterogeneous significant	n=63	Perceived usefulness was the most important factor determining users' intention to use this technology in the future.

2.6 Systematic Review Findings

The systematic review highlighted five themes, with two overall key themes: positive findings of technology use (including nine sub-themes) and negative findings of technology use (including four sub-themes). The other three themes were identified as follows: learning task, learning theories and evaluation method for the technology being used. These themes and sub-themes will now be discussed in turn.

Table 2-6 presents the positive and negative findings and the frequency of these findings; it also shows the breakdown of the sub-themes within both the positive and negative findings themes.

Table 2-6 Positive and Negative Themes that Emerged from the Systematic Review

Positive Sub-themes	**Frequency**
Improved learning	106
Engagement	57
Enjoyment	37
Motivation	29
Collaboration	26
Safe learning	23
Personalised learning	18
Creativity	12
Ease of use	12
Negative Sub-themes	**Frequency**
Technical difficulties	9
Ease of use	4
Pedagogical challenges	2
Irrelevance to learning	2

In relation to the positive implications of technology use, taken together, the papers reported that the use of VW/VR/AR in HE has a net positive impact upon student learning. All 50 papers discussed a total of 22 positive findings, including reduced module failure rate. Improvements in the student experience were frequently reported in the areas of learning abilities/competencies, internalisation of learning, sense of immersion and presence, motivation, engagement and enjoyment of learning.

It was noticeable that three sub-themes' reflected the conclusions of previous researchers, including enhanced learning (Huang et al., 2013, Hwang and Hu, 2013), improved collaboration amongst students (Mikropoulos and Natsis, 2011, Cheong, 2010, Konstantinidis et al., 2010), and authentic learning opportunities (Wu et al., 2013). Figure 2-2 is a summary of the most frequent positive and negative themes drawn from the systematic review.

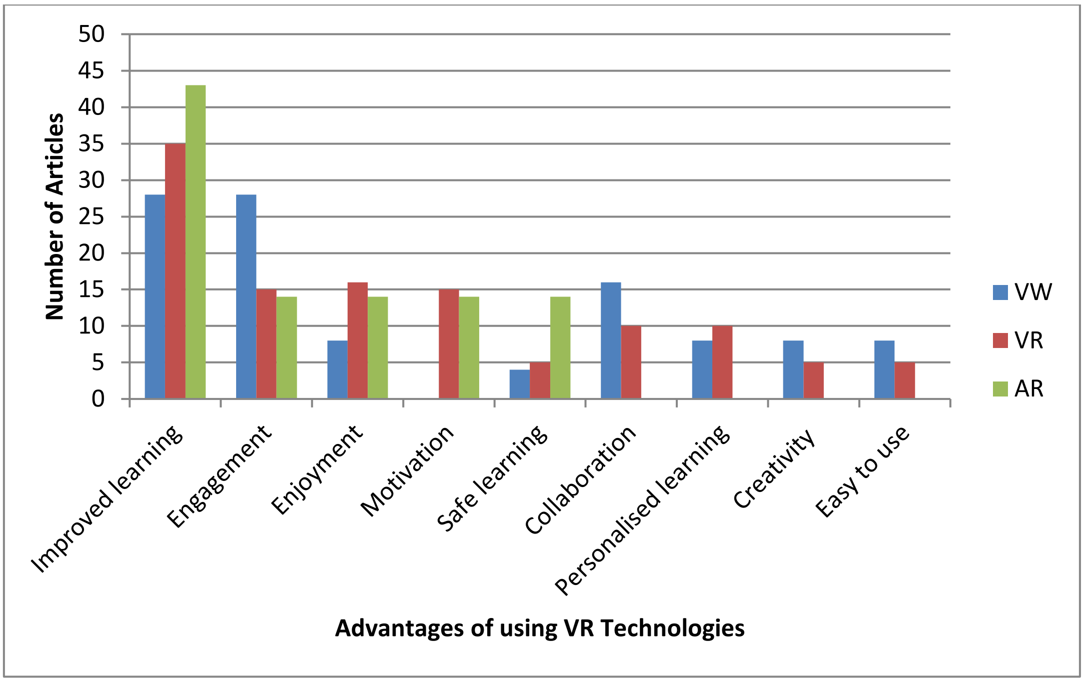


Figure 2-2 Advantages of using VW, VR and AR technologies for Higher Education

VW-based activities were more likely to elicit student immersion and engagement than the other two technologies and provided more opportunities for student collaboration (Figure 2-2). However, the low use of control or pre-test methods in VW papers (discussed below) may skew these results. VR technologies tended to be the most enjoyable and motivating for students (Verkuyl et al., 2020). By comparison, AR studies were found to provide opportunities for improved and safer learning more frequently than the other technologies (e.g., Turkan et al., 2017).

In relation to negative outcomes, four recurrent themes were evident: pedagogical challenges (e.g., Birt et al., 2018), ease of use – e.g., difficulties in using collaborative settings (e.g., Bower et al., 2017), technical difficulties (e.g., Englund et al., 2017; Hack, 2016), and the opinion that the intervention could be a niche application, which might not be relevant to student learning (e.g., Matthew, 2012).

Ease of use was cited both as a positive theme and a negative theme. VR studies cited fewer disadvantages than VW studies and tended to focus on issues surrounding aligning the VR activity with course learning objectives (e.g., Keskitalo et al., 2011). All these themes are supported by previous researchers (Fernandez, 2017). However, Akcayir and Akcayir (2017) stated that they believe such challenges are relatively minor, and that they should not prohibit the use of such technologies. Figure 2-2 suggests how the different characteristics of the three technologies under study might impact on student learning. However, the underlying educational paradigm being tested in the research, and the methods used to carry it out, are likely to impact upon these findings. The systematic review questions addressed these two issues and are discussed in detail below.

In relation to the learning task theme, a range of different practices were reported, including six AR-based studies, and 44 VW/VR-related studies (one was classified as AR and VR). The types of learning tasks reported also varied, and were comparable to Duncan et al.'s (2012) classification model, in that collaborative simulation activities were most cited. This might be explained by the large number of VW studies included in this review.

Underpinning learning theory was another identified theme. An introduction to HE learning theories and approaches was provided in Chapter One of this thesis. In terms of aligned learning theories, the conclusions from this systematic review agreed with those of Mikropoulos and Natsis (2011), that where a pedagogical theoretical model was proposed, it was nearly always based on constructivism, often implied, or was a variant of the approach, and that remaining articles did not refer to a learning theory. This indicates that pedagogical considerations are being overlooked at the design stage of VW/VR/AR learning approaches. Within the broader paradigm of Constructivism, eight variations were identified by Dass et al. (2011) as follows: experiential, contextual, situated, constructionist, transformative, and Connectivist theories, and problem-based and collaborative approaches (Figure 2-3).

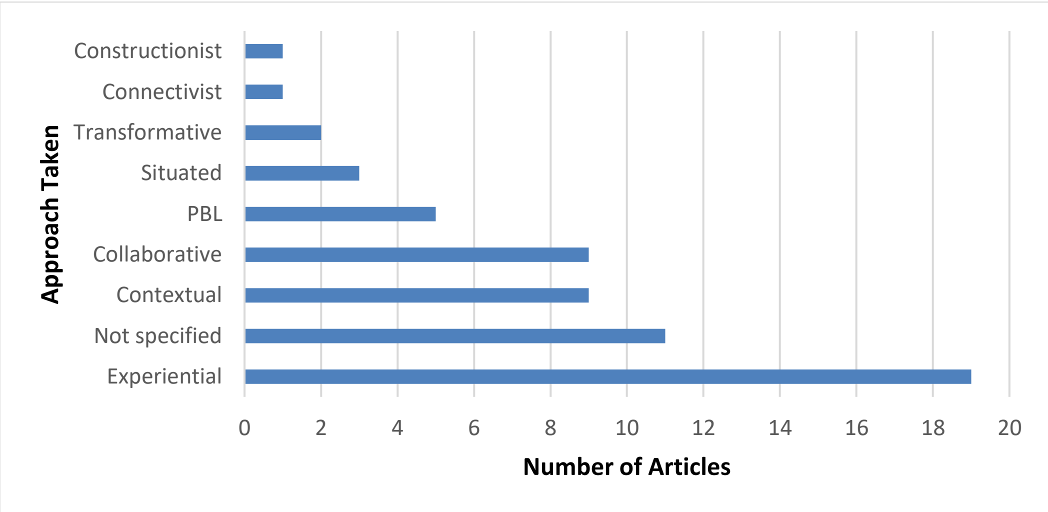


Figure 2-3 The Eight Constructivist Theories/Approaches Reported in the Articles Reviewed

Within the learning theory theme, experiential theory was most frequently cited. This finding dovetails with both Mathews et al. (2012) and Lee et al. (2012) who believe that VWs are experiential by nature. VR and VW enable users to safely "experience" things they might not usually experience. Studies that highlighted the collaborative aspects of using these technologies were those where participants were physically situated in different locations (e.g., Martin et al., (2010). Equally as frequent was the alignment of studies with contextual learning. This, once again, is supported in the literature, for instance by Chen and Tsai (2012), who found that although content is essential in learning, the context with VR learning scenarios is crucial.

Finally, the method of evaluation was a theme that was identified. Analysis of the different technologies in use revealed that whilst VW based studies tended to have larger and more heterogeneous samples when compared to VR and AR, they were less likely to employ pre-testing control and objective measures to evaluate their research (see Figure 2-4). The finding that there were more longitudinal studies conducted for VW might reflect the fact that the technology has been in use for longer.

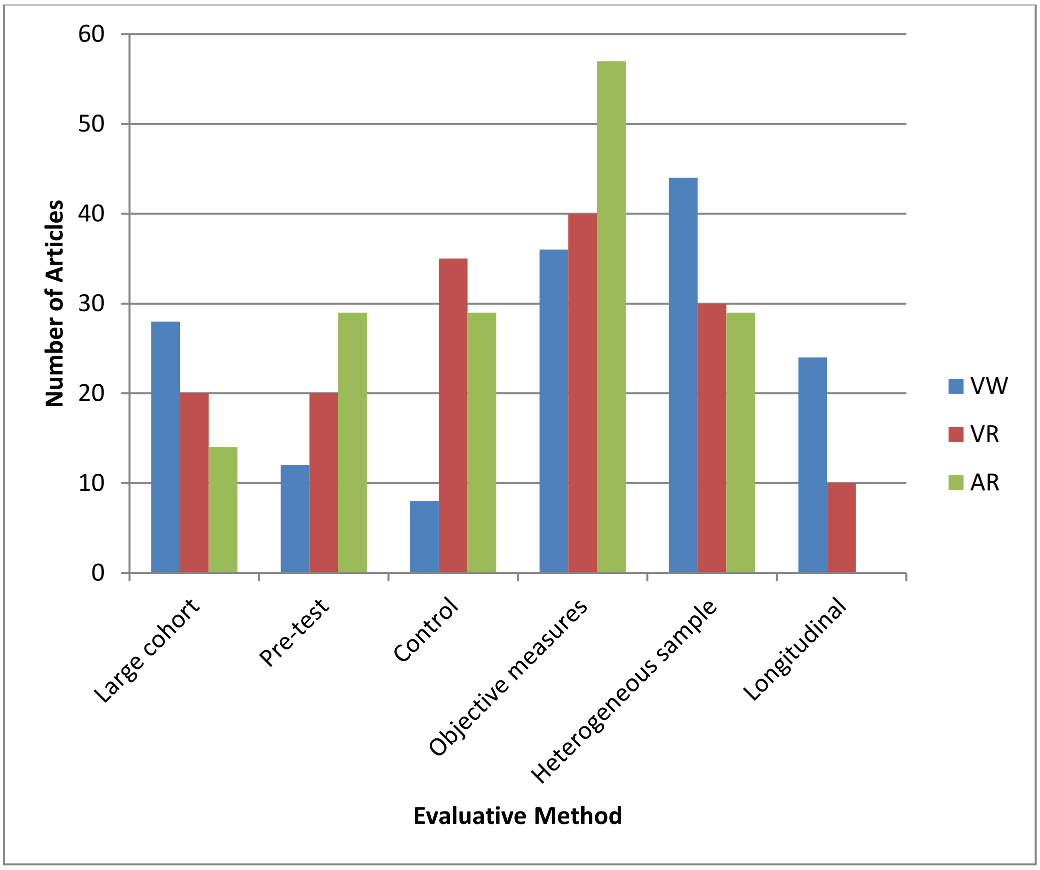


Figure 2-4 Evaluative methods used, broken down into the three categories of VW/VR and AR

Within the method of evaluation theme, studies which employed either a pre-test and or a control were able to evidence that the implementation of the technology (VW/VR or AR) was likely to be responsible for any improvement in the student experience. Figure 2-5 indicates that studies reporting an improvement in student cognition, (mainly VW based as explained above) that have not used a comparison in the form of either a pre-test or a control might be overly positive.

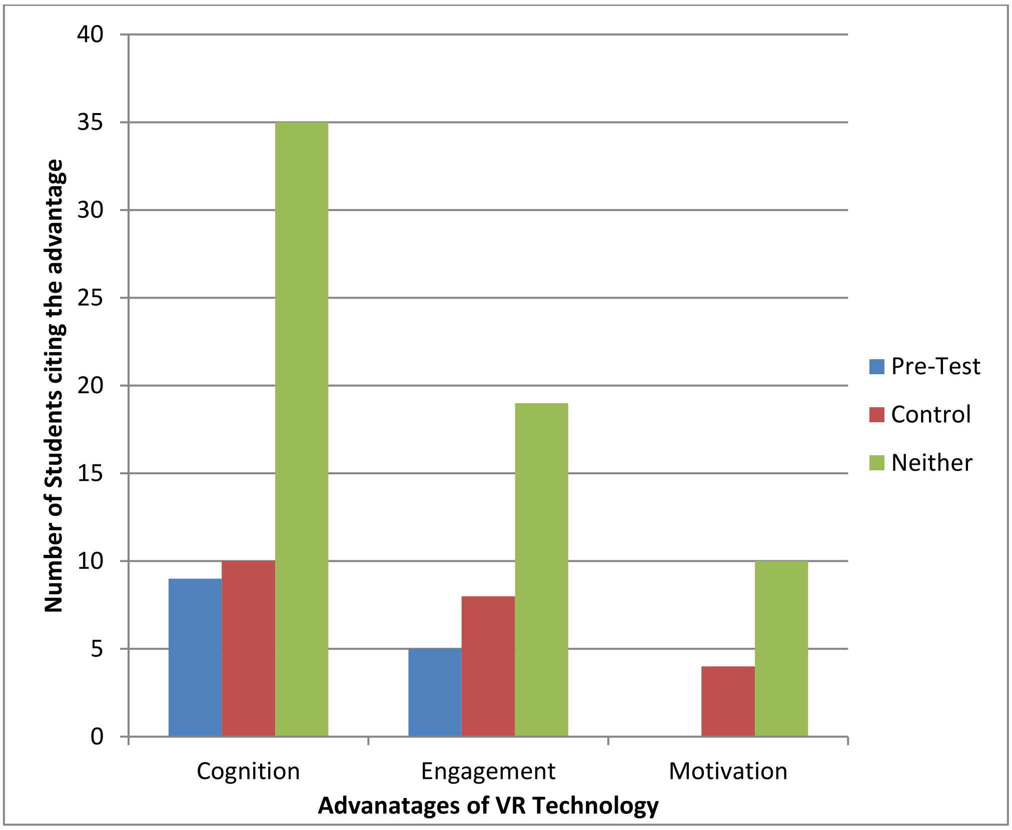


Figure 2-5 Analysis of the advantages of using technology, showing that pre-testing and control groups were not frequently used in studies which reported improved student cognition.

Staying with the method of evaluation theme, the principal findings evidenced that, of the quantitative and mixed methods papers reviewed, most studies could not confirm that improvements in students' learning were attributable to the use of VW/VR/AR, because they had not used a control group, and had not established group equivalence prior to conducting achievement comparison, although it is noticeable that research into the use of VR and AR tended to use pre-test or control methods significantly more than VW research. The reasons for this are not clear from this review, but are worthy of further investigation; it may be that the result of this tendency is that the quantitative evidence underpinning the impact of VW/VR/AR technologies on student learning in HE as a whole could be stronger and more reliable than at present.

Hew and Cheung (2013) found that most studies had relied upon self-reporting to evaluate the efficacy of technology use. However, in our review over half of the studies adopted more objective methods, indicating a recent strengthening in study design and greater attempts at reporting reliable findings. The objective measures reported were wide-ranging and included assessment, skills tests, artefact creation, games trials, and attendance records. Researchers who utilised these objective measures tended to aggregate several distinctive types of evaluation together. Observations were incorporated in just over one quarter of studies, some of which used video capture. None of the studies reported any inter-observer or intra-observer agreement reliability, advocated by Hew and Cheung (2013) to combat the biasing expectations of human observers and inconsistent recording methods due to fatigue.

Overall, studies that employed observations emphasised the collaborative advantages of VW/VR/AR learning. An illustration of this was the evaluation carried out by Wimpenny et al. (2012), who analysed over 130 hours of observational data. They concluded that with sound pedagogical decisions and carefully considered reasons for using VWs, educators can use technology to transform rather than substitute its application. Qualitative studies such as this one appeared to provide more pedagogical direction, suggesting that overcoming the pedagogical challenges of VW/VR/AR technologies requires nuanced evaluation.

Testing was implemented in just under a quarter of studies. For example, Hack (2016), used evaluation of student grade and attendance and produced very precise conclusions: “module failure rate was less than 5% compared to 11% in the previous two cohorts”. However, Hack’s (2016) study failed to employ a control or pre-test, therefore it is difficult to be certain that the student grade improvement was related to the technological intervention.

Although the inherently subjective nature of self-reporting methods deems them unreliable (Cohen et al., 2017), some researchers made efforts to fully describe the methodology behind their questionnaire and its relative reliability (Liu and Chu, 2010). Reliability was also considered for interview design with some researchers using a range of different types of interview including pre/during/post-intervention iterative interviews (Marcel, 2019). However, the studies that employed interviews failed to report measures of authenticity such as member checking, which Hew and Cheung (2013) deem to be good practice. Academics need reliable and valid findings on which to base important decisions when selecting technologies to support curriculum objectives.

In relation to the theme of evaluation method, returning to Kurilovas’ (2016) finding that most innovations were evaluated by small-scale homogeneous case studies, this evidence supports this finding, as it was discovered that studies continue to use inappropriate sample sizes and inappropriate sampling strategies. However, one finding was that there has been an increase in the use of pre-testing over the last four years (before 2013 n=1, post 2015 n=10). Ghanbarzadeh and Ghapanchi’s (2018) appraisal identified no longitudinal studies, and whilst a fifth of the papers in this review were classified as longitudinal, more in-depth investigation of the retention of learning over time requires the implementation of further longitudinal studies.

Having discussed the findings thematically, limitations will now be identified. The first limitation is that only peer-reviewed published papers were included and therefore publication bias might have affected the findings. The decision to exclude papers published in other languages is a further limitation, though was unavoidable. Finally, only seven AR studies were reviewed, and this limits the generalisability of findings related to AR. However, the search strategy was sufficiently thorough, so that the review included all the relevant papers.

2.7 Conclusions of the Systematic Literature Review

Two questions guided the systematic review, which were:

What underpinning learning theories and practices have been reported for VW/VR/AR learning systems in HE?

How has the effectiveness of VW/VR/AR learning systems been evaluated?

In relation to the first question, the types of learning tasks reported varied, with collaborative simulation activities most commonly cited. In terms of aligned learning theories, constructivism (and related theories) remains dominant. VR and VW were found to enable users to safely “experience” things they might not usually experience (experiential learning theory). The context was found to be a vital deliberation for VR learning scenarios. Overall, the review has shown that pedagogy requires more in-depth consideration when designing VW/VR/AR learning approaches.

In relation to the second question, this systematic review has identified that self-reporting is no longer a preferred method of data capture, and that researchers are using a variety of objective evaluative measures including observation and assessment, as well as video analysis of intervention sessions. The most rigorous studies combined objective evaluative measures, pre-testing and use of a control. Future researchers should consider which methods are suitable for their investigations and combine them for optimal impact. The power of qualitative research techniques (including subjective techniques when appropriate) is also recognised, particularly those techniques that can add depth (Girvan et al., 2013).

This study argues that whilst measurements have become more objective, stronger evaluation is necessary to confirm the efficacy of VW/VR/AR use in HE, particularly in relation to the use of pre-testing or control group comparisons in VW research. This will be even more important when AR begins to be used more widely, as clearer proof of efficacy is likely to improve scalability of these technologies and convince academics that investing their time and faculty money in these technologies will be justifiable. Longitudinal, large-scale, experimental studies are also required to evaluate these technologies more holistically.

This systematic review concludes that the implications for Practice and Policy are that educators can use this technology to enhance students' internalisation of learning in HE, but that further studies need to evaluate cost implications of this technology. By reviewing the recent studies on the use of VW/VR/AR in HE and adding new findings, it is hoped that this study can provide policymakers and practitioners in the education sector with new insights into effective methods to evaluate the application of these technologies.

What this systematic review adds:

- Insight into current VR/VW/AR practices in HE, their underpinning theories and pedagogies
- Critical analysis of the various methods that have been used to evaluate the effectiveness of VR/VW/AR technologies in HE

2.8 Chapter Two Summary

Within this second chapter, the literature surrounding VW/VR/AR use in HE was reported and explored as a systematic review. The review signposted that investigators found the use of such technologies to be beneficial; however, numerous cautions were made, e.g., ensuring activity relevance, providing students with the opportunity to learn how to use the software, finally making sure that students have access to the required technology to run the software. Through building upon the evidence from this review, the evaluation of the teaching and learning intervention was designed to be as rigorous as possible; specifically, the empirical study was an RCT. Having based the empirical research design upon the outcomes of the systematic review, a second literature review was undertaken to identify conceptual considerations and to develop general and specific hypotheses.

Having undertaken a systematic review in order to provide direction in terms of how the evaluation of a VR intervention for nurse education can be robustly conducted (e.g., via using objective evaluative measures, pre-testing and use of a control), the next step (Chapter Three) was to conduct a CR review in order to identify the different mechanisms (or variables) that might be interactive when student nurses are educated using VR technology in a large-scale lecture theatre setting.

Chapter 3 Critical Realist Review of the Literature

3.1 Synopsis

The systematic literature review reported in Chapter 2, focused on guidance concerning the suitability of VR in the HE context being addressed and the optimal evaluation of an RCT. However, a broader review was required to identify factors (latent variables) that might come into play during the use of VR as a teaching tool in an HE lecture scenario. A CR review concentrates on reviewing literature related to multifaceted social interventions, processes, and practices (Edgley et al., 2016), hence its suitability for this research.

A CR review tries to unpack the compound social phenomena in terms of their constituents, the interfaces between components with the context, and the mechanisms involved in causal outcomes. This approach covers the complexity of social phenomena in complete breadth and depth (Allana and Clark, 2018). A CR review approach is suitable for evaluating interventions that are composed of several interacting components which are affected by contextual factors. Thus, CR reviews of such interventions attempt to unravel: *What works the best for whom, where, how, and why?*

This chapter reports the findings of the CR literature review and the findings fed into the conceptual framework tested as part of the RCT. CR philosophy is explored including which approach towards CR has been applied within this thesis. Key terms specific to this chapter, including Immersion and Critical Realism (CR), are briefly discussed, before the literature review is reported. The chapter closes by setting out the proposed conceptual framework for the action of VR learning within an HE setting.

In Section 1.3 a brief overview of presence and immersion was presented; the terms will now be explained further. In the literature, the terms presence, flow, and immersion are often used as having the same meaning and origin (Klevjer, 2008, Childs, 2010, Dalgarno and Lee, 2010, Mikropoulos and Natsis, 2011). They are sometimes referred to as subcategories of “user experience” (Dalgarno and Lee, 2010). Witmer and Singer (1998) stated that a Virtual Environment: “that produces a greater sense of Immersion will produce higher levels of presence.” This assumes that immersion is a precursor for presence. Michailidis et al. (2018) concluded that the terms can be used interchangeably: “an immersive virtual environment is one that perceptually surrounds the user and could increase their sense of presence.”

Similarly, flow and immersion definitions overlap (Kiili et al., 2012, Radianti et al., 2020). For instance, attentiveness, loss of time perception, a balance between the player’s skills and the game’s demands, and loss of self-awareness are some of the shared assets that both flow and immersion exhibit (Lazaros et al., 2018). Flow has been referred to as the optimal experience when nothing else matters (Lazaros et al., 2018) being more intense and possible extreme than immersion, which infers that immersion is sub-optimal in some ways (and therefore, more suitable for video gaming).

Three grades of immersion: engagement, engrossment and total immersion, were discussed by Lazaros et al., (2018), who went onto describe how the model of an average immersive experience in video game playing could be viewed as being reduced to the engagement and engrossment levels, whose characteristics are not considerably divergent from flow. Lazaros et al., (2018) concluded that given the similarities between flow and immersion, it is not safe to conclude that flow is more “extreme” than immersion, or that they are conceptually different. Their proposed variances are not persuasive enough to set immersion apart as a dissimilar mental state. Indeed, Taylor and Dunne (2011) stated that presence “*forms the very foundation on which*

Immersion is built". Hence in this thesis, whilst it is acknowledged that presence, flow, and immersion are distinct, it is thought that an overarching theme of "immersion", which encompasses both concepts of flow theory and presence, is sensitive enough for the research that has been conducted, particularly in the light of literature which groups the themes as overlapping and/or interchangeable.

3.2 Philosophy of Critical Realism (CR)

CR is an important perspective in contemporary philosophy and social science (Marchal et al., 2012, Dalkin et al., 2015) and in the field of healthcare research (Linsley et al., 2015). Realistic evaluation is a form of evaluation that is determined by theory and was developed by Pawson and Tilley (1997a), founded on the philosophy of CR (Bhaskar, 1989). Realistic research strives for a clear understanding of interrelated structures and measures as well as the contextualised social situation of the explored subject matter.

Central to CR is a rejection of the assumption that the effectiveness of an intervention is based only on its inherent qualities. CR instead proposes that outcomes result from complex interactions of causal mechanisms, which differ according to context (Spacey et al., 2019). Mechanisms are embedded in both the intervention itself and in the social/organisational context in which the intervention is introduced. Furthermore, such mechanisms are filtered through the eyes of humans. Therefore, evaluation of an intervention's efficacy must include how different people experience and respond to it and for what reasons. Traditional experiments ask: 'Does this work?' or 'What works?' CR evaluation adds to this by examining: 'What works for whom in what circumstances?' (Pawson and Tilley, 1997a). This means that outcomes (of interventions) should move beyond simply measuring and recording, towards studying the implications of interventions for human lives.

Three of the main features of Bhaskar's (1989) CR philosophy, which underpin the research performed in this thesis, are the importance involved in (1) generative mechanisms, (2) the stratified character of the real world, and (3) the dialectical interplay between social structures and human agency. These features will now be discussed in turn.

The first main feature of CR is the importance of generative mechanisms. Critical realists uphold that generative mechanisms, though not directly observable, are nonetheless real and can be recognised through their effects. The CR view of causation is essentially different from the concept of causal laws linked with positivism, in which causation is seen as being associated with the continual combination of events connected in time and space (Brant, 2001). Critical realists assert that the world works as an open system with many dimensions (Benton and Craib, 2005), and that generative mechanisms may stay latent until triggered in specific circumstances.

The second key idea central to CR is that of a stratified reality. CR recognises that the social and natural worlds are characteristically stratified, with causal mechanisms operating at various levels of reality (McEvoy and Richards, 2006). For example, a mental illness might emerge as the result of interactions between different levels of stratified reality. The trajectory of any mental illness might be influenced by societal interactions as well as what is going on in an individual’s mind and body for instance.

Finally, the third idea is that critical realists have emphasised the interdependence of structure and agency. Social structures afford resources that enable individuals to act, as well as providing boundaries for individual behaviour (McEvoy and Richards, 2006). However, the behaviour of human agents is not completely determined by social structures. This is because agents are also able to alter social structures by reacting to different circumstances. Drawing upon the work of Bhaskar (1989), Kerr (2017) defined mechanisms as “the unseen ensemble of tendencies, liabilities, and powers possessed by objects in the world.” Human agency is the principle amongst this ensemble (Higgins et al., 2012).

Archer et al.’s (2013) conceptualisation of human agency has shaped this thesis, upholding that human agency is the ability to think, reason, imagine, plan and believe. Archer et al. (2013) provide a framework which demonstrates interrelationships between agency, structure and culture. Differing from structural theory, where agency and structure are regarded as ontologically and analytically inseparable, CR philosophy perceives them as being closely intertwined. It also diverges from the discourse’s analytical slant that asserts it is only through the discourses and practices that social dealings are represented (Hjørland and Wikgren, 2005).

Human agency needs to be treated separately from social mechanisms. An example of research conducted that highlighted the importance of handling agency distinctly from social mechanisms was confirmed in a realist evaluation by McConnell and Porter (2017). Different agents (e.g., nurses and consultants) have tended to interpret changes in objective conditions that resulted from the withdrawal of resources in diverse ways. There have been differences in the way nurses and consultants responded. While being challenged with the same social mechanisms in relation to the removal of resources, these different groups of agents tended to preserve and/or alter their social context in dissimilar ways, depending on their interpretations.

Moving on to examine outcomes in relation to mechanisms, researchers, including Pawson and Tilley (1997), Jagosh et al. (2014) and Linsley et al.

(2015), believe that causal outcomes follow from mechanisms working in contexts. Their work has evaluated causal mechanisms (M) and the conditions (C) under which they are triggered to produce specific outcomes (O). Hence, they have used context-mechanism-outcome (CMO) configurations to evaluate interventions.

According to Pawson and Tilley (1997), a realist approach recommends that: the *outcomes* (consequences) of an intervention are subject to the interface between *mechanisms* (i.e. the provision of intervention *resources* and concepts and how the implementer and recipient respond to these through *reasoning*) and the *context* (the environmental background). This process is abbreviated to CMO. Each CMO forms the foundation of a 'mini experiment'. By evaluating a sequence of CMOs, it is conceivable to infer the topographies of contexts that permit diverse mechanisms to act to attain outcomes. Consequently 'transferable lessons' may be learned (Pawson and Tilley, 1997). One oversight in the work of Pawson and Tilley (1997) is that it does not account for the dialectical interplay between social structures and human agency that is considered to be a fundamental element of CR philosophy (Archer, 2003; Bhaskar, 1989; Danermark, 2002; Higgins et al., 2012; McConnell et al., 2015; McEvoy and Richards, 2006; Porter, 2015; Pratten, 2020; Rutzou, 2017; Williams, 2020).

Drawing upon Bhaskar (1989), and Archer (2003) and rejecting some of Pawson and Tilley's (1997) theories, Porter (2015) defines 'contextual mechanisms' as the conventional mechanisms faced by programme initiators, and 'programme mechanisms' as the variety of mechanisms enclosed in programmes intended to countervail against the influences of contextual mechanisms. Therefore, Porter (2015) proposed adaptions to the CMO configuration equation to include CM + PM = O. Programme mechanisms have also been evaluated by other recent CR researchers (Eastwood et al., 2014). These adaptations (Table 3-1) are also in agreement with Archer's morphogenetic method (Archer, 2013), which includes contextual mechanisms and programme mechanisms within the sequence.

The distinction of agency from social mechanisms described in an earlier section, (e.g., Archer, 2003; Bhaskar, 1989; Higgins et al., 2012; McConnell et al., 2015) was also highlighted by Porter (2015), who recommended that the CMO configuration approach should also include the component of ‘A’ for Agency in order that the equation provides a satisfactory report of social causation. Hence, the reviewed evaluation formula takes the form of Contextual Mechanisms + Programme Mechanisms + Agency = Outcome (CPMAO configurations), for which the equation is: CM + PM + A = O. This modified equation forms the CR approach that this thesis will take.

3.3 Examples of CR in Nursing Research

The value of CR has thus far been established, mainly through the social science disciplines including research conducted by Alderson (2020). However, a substantial amount of CR literature has materialised from researchers in various other disciplines including economics (Westra, 2019), management (Thorpe, 2020) and marketing (Simmonds, 2018).

There are also examples of where CR has been used for nursing research (Spacey et al., 2019, Williams et al., 2017, Ryan, 2019, Coleman, 2019, Aspinall et al., 2019, Schiller, 2016). Education and health care both comprise “complex human interactions that can rarely be studied or explained in simple terms” (Coleman, 2019) and there has been substantial debate in these arenas regarding the most suitable research methods to best enlighten policy and practice. This is one reason for more holistic methods, such as CR being used to underpin research. CR can be extremely useful in exploring complex interventions holistically, including their components, contexts, and mechanisms (Allana and Clarke, 2018).

Bellass et al. (2019) adopted a realist review method to detail the key findings from the body of work that links the experience of living with dementia to the neighbourhood. The use of the CR method enabled them to pinpoint three mechanisms that interact, namely: outdoor spaces, built environment, and everyday technologies. It also allowed insight into relationships and interactions (between mechanisms) that require further research, e.g., the neighbourhood as a social space and as a physical space alongside the active role of people with dementia as 'place-makers'.

Porter et al. (2017) examined the processes and experiences involved in the introduction of music therapy to palliative care in the UK. Using a realistic evaluation approach, they conducted a qualitative study (using surveys and focus groups). Through their CR approach they were able to identify that music therapy contains multiple mechanisms that can provide physical, psychological, emotional, expressive, existential and social support.

A systematic CR review was conducted by Spacey et al. (2019), with the aim of describing and explaining the effectiveness of interventions designed to improve end-of-life care in care homes; 41 studies were included in the review. The method enabled the following nuanced findings: there was evidence to suggest that education and inter-professional collaboration can be effective intervention mechanisms for improving end-of-life care, and high staff turnover was a significant contextual mechanism impacting on the sustainability of interventions. In terms of human agency, the dedication and enthusiasm of care home staff who deliver end-of-life care was highlighted.

Together these studies demonstrate that CR can be used as a more holistic method to inform policy and practice, through identifying not only the agency and mechanisms involved in complex health interventions but also more subtle interactions and relationships between these mechanisms.

3.4 Critical Realist Literature Review Process

According to Edgley et al. (2016) a CR review is an appropriate method for researchers who want to look at *why* they see the things they do in practice. Edgley et al. (2016) believe that the CR review is an organic process, saying that it is:

> *"an intellectual and personal journey in which the pathway is created as a direct result of the researcher's critical choices, while the destination similarly cannot be known until it is reached. Critical realism offers sound methodological justification for combining empirically-based descriptions of what is, exploration of the 'real' and reasoned conclusions about how organisations and practices should be."*

In this respect, research practitioners are encouraged to adopt a theoretical viewpoint, influenced by their values on a subject matter. A CR review methodology was therefore chosen as it helps to discover new ideas, potential new theories and allows the beginning of a critical inquiry (Edgley et al., 2016) without quite knowing where it might end.

The objective of this CR Literature Review was to develop theory that explains the underlying mechanisms that produced observed outcomes when using VR for learning in an HE setting. Whilst VR is generally thought to be beneficial to student learning, what is still not clear is how VR applies its effect, for whom it works best, and in what situations. There is a growing request, not only for evidence-based practice but also for in-depth appreciation of how an intervention works in order to improve attainment (Pawson, 2012). Consequently, the primary questions for this review were: What seemed to be the learning mechanisms? For whom did they work? (and) Which contexts influenced the intervention to accomplish its anticipated outcomes? (McConnell and Porter, 2017). Hence, the overall question addressed by the realist review, designed to provide contextual indicators to support explanation of the observed phenomena, was:

"What factors influence student learning when using virtual reality?"

For a Realist Synthesis process, following sets of procedures will not ensure that a review will be robust (Jagosh et al., 2014). Rather, it requires a series of judgements about the significance and robustness of specific data for the purposes of answering a specific question. Data are selected on the grounds of two criteria (Jagosh et al., 2014): relevance (can it contribute to theory testing/building?) and rigour (were credible and trustworthy methods adhered to in order to generate the data?). Repeating patterns of CPMAO configurations are likely to be found across the included articles/documents.

To develop the theoretical framework on VR for HE students, a broad and rapid review was conducted. The aim was to build up a representative body of literature to aid testing and refinement of the intervention theories, not to comprehensively search all available evidence (Jagosh et al., 2014). The literature search was iterative and continuing. This decision was based upon the idea that as the synthesis progresses, new or polished features of theory may be required to explain findings, or to scrutinise precise facets of specific measures.

As new elements of theory were included, searches for evidence to support, refute, or refine those features were incorporated. Finally, as a last but potentially fruitful step, the technique of examining the reference lists of previous, related literature reviews was adopted. A wide range of cross-disciplinary subject databases was employed because rich explanations are the goal of CR reviews. Searched electronic databases included: ERIC, Education Source and CINAHL from their first available date until April of 2019 using “virtual reality,” “Higher Education,” and “learning outcomes” as broad search terms. After the initial search, 81 articles were retrieved (Figure 3-1).

virtual reality AND higher education AND learning outcomes	**Expanders** - Apply equivalent subjects **Search modes** - Boolean/Phrase	**View Results** (81)

Table 3-1 The Critical Realist Review Search String

Following the realist method, all pertinent literature was examined to advance theories that might explain how VR learning works. Additional studies were located through pearling (examining the reference lists of included studies). Searches were limited to publications in the English language. Title and abstract were screened first; potentially relevant literature was then looked for in full text articles. Articles were incorporated if they met the inclusion criteria for each phase of the searching process (they first needed to address VR for an HE population, and for subsequent searches they needed to address developing theoretical assumptions). Studies were excluded if they involved schools. Codes were shaped as the synthesis advanced to portray new data for concept analysis, and articles were revisited to safeguard that coding was complete and consistency preserved. Predominantly recurrent themes and enabling contexts were recorded to help identify the underlying programme mechanisms. Iterative searching provided auxiliary data to cultivate and test theories.

When extracting explanations, from a CR perspective, these explanations should be offered both in terms of reasons for relationships between concepts (horizontal explanation), and in terms of fundamental structural mechanisms that generate events and that bring about observed phenomena (vertical explanation). From a CR stance, both quantitative and qualitative studies are valued; hence data were extracted from both types.

3.5 Findings of the Critical Realist Review

Thirty-six papers were retrieved from the CR review of the literature surrounding factors that influence HE students when learning using VR. Overall, it was found that research on the use of VR technology in education is in its embryonic stage. Both quantitative and qualitative studies have reported generally positive results for learning with VR (Mason and Holmes, 2018; Slavova and Mu, 2018).

An example of a study with positive findings is that of Dubovi et al. (2017), who aimed to appraise the success of desktop VR when applied to nursing education, as a tool for learning medication administration measures. A quasi-experimental pre-test-intervention-post-test-comparison group strategy was conducted based on quantitative analysis of surveys, video recordings and worksheets. An experimental group (n=82) was compared with a lecture-based curriculum group (n=47). The outcomes exposed significantly greater theoretical and procedural knowledge learning gains following activity with the desktop VR simulation compared to the control. Desktop VR exposed the students to their own errors, permitting learning practice followed by continuous feedback essential to skill attainment. Although the simulation was constructed on a desktop VR, it enabled a robust sense of presence. A slight positive association was found between the sense of presence, predominantly the sense of control, and conceptual learning of medication administration. This shows that by cultivating students' sense of control in the desktop VR, the learning process can be enhanced. Thus, VR simulations may deliver inexpensive and malleable opportunities to practise the required practical skills in HE, which is vital to developing learners' proficiency.

However, there were also some studies that found that whilst VR produces more Immersion, it can negatively impact upon HE students' learning (Guido et al., 2019). Makransky et al. (2019) tested a sample of 52 university students who partook in an experiment in which students learned from a science simulation via a desktop display (PC) or an HMD (VR); and the simulations enclosed on-screen text or on-screen text with narration. Their discoveries were that students described being more present in the VR condition; but they learned a smaller amount and had a significantly greater cognitive load based on the electroencephalography (EEG) measure. Despite its inspiring assets (as reflected in presence ratings), learning science in VR may overload and distract the student (as reflected in EEG measures of cognitive load), ensuring less chance to build learning effects (as reflected in inferior learning outcome test performance). Cobbett and Seagrove-Clarke (2016) found no significant difference between VR and traditional learning methods in their RCT. However, the sample size of 56 was a limiting factor in their research.

Four main themes emerged during the review process, which led to the development of the four separate but interrelated latent variables (or mechanisms) that encapsulate VR as an approach to teach students in HE and influence their learning outcomes or short-term knowledge gain. The four main themes were as follows:

- student CONFIDENCE (six papers retrieved)
- students' prior EXPERIENCE (four papers retrieved relating to VR and two more general papers retrieved)
- student ENGAGEMENT (nine papers retrieved)

- student IMMERSION (22 papers retrieved)

Table 3-2 lists all papers that were identified as being relevant during the CR Literature Review. Some articles were categorised in more than one theme.

Table 3-2 Articles retrieved from the Critical Realist Review of the Literature Organised Thematically

Identified Themes	Definition of Theme within the context	Paper author/s and date
CONFIDENCE	Student confidence in their own learning/skills	Cliffe (2017); Cobbett and Snelgrove-Clarke (2016); Parong and Mayer (2018); Schunk (1995); Smith and Klumper (2018); Su and Cheng (2013)
EXPERIENCE	Computing experience and/or past subject matter related experience (e.g., experience of looking after a patient with diabetes)	Childs (2010); Falconer (2013); Hill (2017); Liaw et al. (2019); Michailidis et al. (2018)
ENGAGEMENT	Engagement with the learning content/activity	Arbaugh and Duray, 2002; Childs, 2010; Klevjer, 2008; Lee et al., 2010; Mayer et al., 2008; Parong and Mayer, 2018; Pintrich, 2002; Slavova and Mu, 2018; Sun et al., 2008
IMMERSION	Immersion in the VR learning; this could be influenced by the student's immersive tendencies as well as the fidelity of the software, and so forth	Bailey and Witmer (1994); Bulu (2012); Csikszentmihalyi (2000); Dubovi et al. (2017); Falconer et al. (2018); Huang et al. (2013); Janssen et al. (2016); Johns et al. (2000); Liaw et al. (2019); Makransky et al. (2019); Mason and Holmes (2018); Milk (2015); Nunez and Blake, (2001); Singer et al. (1997); Slater and Wilbur (1997); Stevens and Kincaid (2015); Tcha-Tokey et al. (2018); Tüzün and Özdinç (2016); Warburton (2009); Witmer and Singer (1998); Wong and Lee (2015)

The literature will now be discussed using these thematic headings.

3.5.1 Confidence Theme

The first theme was confidence. Literature examining the impact of confidence on VR learning was sparse. Six studies clearly identified a link between VR learning and improved student confidence. For example, Cliffe (2017) found that VR field trips were beneficial in constructing student skills and confidence in an organised environment pre field trip. Su and Cheng (2013) demonstrated improved confidence when learning engineering with VR compared to traditional methods via their RCT. Similarly, Smith and Klumper (2018) found improved confidence after VR training in their pre-test/post-test study; albeit, a sample size of 15 participants limits these findings.

One study, however, did not find confidence enhancement following VR training. Cobbett and Snelgrove-Clarke (2016) found no significant difference between face-to-face learning and VR clinical simulation in terms of student self-confidence when learning about maternal and new-born nursing concepts. Their study was an RCT with pre- and post-testing and included 56 participants.

An identified sub-theme of confidence was self-efficacy. According to Parong and Mayer (2018), the case for using immersive VR for educational purposes is rooted in interest theory and self-efficacy theory. Interest theory concerns motivation and will thus be discussed in Section 3.6. Self-efficacy philosophy states that students work harder when they perceive themselves as competent for the task; again this is closely linked to a learner’s self-confidence (Schunk, 1995). Self-efficacy, as defined by Schunk (1995), is a person’s judgement of their aptitude to perform an agreed action. Schunk (1995) defined the procedure of self-efficacy and attainment as a feedback circle. Primarily, the student has their principles about their self-efficacy. This self-efficacy then affects the student's task commitment for which they receive feedback. Lastly, this "aptitude" feedback restructures learners' self-efficacy. VR games that integrate a feedback structure for progress on academic content would be acting within this system. For example, various interactions within a lesson could provide suitable feedback that enhances the student's self-efficacy, and hence a student's motivation for the lesson. This may be a benefit over the feedback perceived in old-style lessons as learners would have immediate updates on their self-efficacy.

Overall, for this present study it is anticipated that student confidence would improve in those who use the VR training simulation. Students completing a pre and post-test survey for the current study were thus asked to rate their confidence before and after the training. The feedback loop that might be involved in boosting confidence will also be investigated during this research.

3.5.2 Experience Theme

Experience was the second identified theme. Five studies that were retrieved had examined the impact of experience on learning with VR. One aspect of experience was identified as "prior exposure". Childs (2010) found that preceding exposure to technology did not predispose students to the use of VWs. He established that some students who formerly were not concerned with technology and had little exposure to different technologies still enjoyed the VR sessions. However, in his study, learners with previous exposure to games all felt presence and enjoyed the VR sessions. Notably, transferability of skills in navigation did not seem to be pertinent; gamers struggled just as much with the navigation as did individuals with limited or no familiarity with games (Childs, 2010). Hence it is hypothesised that prior computing experience will not influence student's learning outcome or positive experiences when learning with VR in the research reported in this thesis.

Additionally, there is literature surrounding experience and learning that is not directly about VR learning but is relevant. According to some researchers (Hattie, 2015, Hill, 2017), students' prior experience and knowledge have an impact on how they engage with new learning. Hill (2017) suggested that a student's level of study, e.g. A level, graduate, and so forth, impacts on their approach to learning in the future. For this reason, the student's level of study in education was included as one of the questions in the survey of the current research. For example, if a student had previously completed an A Level in biology, then they might find it easy to understand complex physiological concepts relating to diabetes, compared to students who had not studied biology at A Level. Hattie (2015) found that a student's "prior knowledge" impacted their learning ability and this supports Hill's (2017) finding. For this reason, a pre-test multiple choice question (MCQ) quiz was included in the RCT process in order to ascertain students' prior knowledge in this current study.

Hattie (2015) believes that a student's prior knowledge is part of their "agency". Relating learning attributes to agency, Hattie (2015) considers that students bring dissimilar qualities and prior knowledge, they have dissimilar motivations and purposes for learning, they study in diverse ways, some are collaborators some are loners, they have an assortment of likes and dislikes, and they can be bright or struggling. This links to the notion of agency and the personalised and individualised nature of learning material and the CR approach taken in this thesis.

3.5.3 Engagement Theme

Turning to the third theme, engagement, nine researchers had pinpointed the relationship between use of VR in education and student engagement (Arbaugh and Duray, 2002; Childs, 2010; Klevjer, 2008; Lee et al., 2010; Mayer et al., 2008; Parong and Mayer, 2018; Pintrich, 2002; Slavova and

Mu, 2018; Sun et al., 2008), concluding that interest/engagement in the learning material leads to better student motivation. The link between motivation and engagement is unclear; therefore, the terms will be used interchangeably. Klevjer (2008) stated that engagement “involves those portions of a text where extra effort or interpretive skills are called for, where external referents are sought”. Klevjer (2008) therefore considers both immersion and engagement as being significant parts of interaction with games and states that they are reciprocally dependent. It is when a user moves continuously between these two states that computer simulations or games are at their most enthralling and contribute to the experience of “flow” (Klevjer, 2008). For this researcher, the notion of both immersion and engagement are tightly entwined.

A sub-theme of engagement appears to have been “interest theory”, which was raised by Parong and Mayer (2018), Pintrich (2002) and Mayer et al. (2008). According to interest theory, students work harder when they value the material, with either individual or situational interest (Parong and Mayer, 2018). Individual motivation might not be uniquely primed TEL, as a student’s inherent interest may not be affected by the learning media. However, a novel, immersive technology, such as a VR lesson, could prime a student’s situational interest more than traditional teaching methods (Parong and Mayer, 2018). They stated that student motivation plays a great role in deep learning; those who are more motivated are expected to engage in the lesson or task, put in more effort, and be resilient when overcoming difficulties in understanding.

In relation to interest theory, Pintrich (2002) found that in order to increase the student’s motivation, initially, the lesson may engage student curiosity; then, the learner’s interaction with the lesson may prime their self-efficacy to carry on with the lesson. In the case for VR, the motivating, immersive experience may spur the learner’s individual interest and the feedback from interacting with the lesson should keep them feeling capable to progress their learning. However, other research demonstrates that adding stimulating but extraneous material to a multimedia lesson may negatively impact on learning (Mayer et al., 2008), by distracting them from the intended learning. Research surrounding engagement and VR learning has been inconclusive and warrants further investigation.

Linked with the theme of engagement, Lee et al. (2010) found that usability of VR software was a significant precursor to motivation/engagement in the desktop, VR-based learning environment. They found that learning activities that are regarded as being beneficial and easy to use in a desktop, VR-based learning environment help to motivate students, which in turn affects the learning outcomes. In their study, VR features were indirectly associated with reflective thinking, which was mediated by usability. Lee et al. (2010) concluded that the learning activities and tasks delivered must be worthwhile and easy to use for the desktop VR to entirely fulfil its abilities and capacities to advance learners' learning experience. In short, the design aspects of the desktop, VR-based learning environment that considers the professed worth and ease of use has a substantial bearing on learning experience and learning outcomes. This research echoes that of Lim and Clark (2010) and Arbaugh and Duray (2002) in technology-mediated learning.

Conversely, Childs' (2010) research concluded that the students who appreciated the sessions were no more likely to be able to navigate the VW than those who did not appreciate the sessions. For this reason, supposed usefulness and ease of use questions were included as part of the "engagement variable" survey questions.

3.5.4 Immersion Theme

Moving now to the immersion theme, in total 21 papers were identified that particularly focused on the importance of immersion for VR learning; this was an indication of the significance of the part that immersion might play in the process of VR learning (Falconer, 2017). In Section 1.10, it was stated that for the purposes of this thesis presence, immersion, and flow would be used interchangeably, as the connections between flow, immersion, engagement and presence is indistinct. For Klevjer (2008), disturbance gives rise to engagement, and it is oscillation between engagement and immersion that gives rise to flow. Moreover, numerous researches investigated engagement as being the first state of immersion (Huang et al., 2013). Those students who exhibit an inclination towards mediated presence and identify with and develop their avatar are more engaged and motivated to take part in VW activities, and are more likely to enjoy their learning (Klevjer, 2008).

Immersion has also been found to lead to increased conceptual learning (Lau and Lee, 2015, Witmer and Singer, 1998, Stevens and Kincaid, 2015, Tüzün and Özdinç, 2016). Janssen et al. (2016) found that a “sense of being in the environment” might help some students learn more intensely than students who learn by partaking in the learning task as observers. Thus, a constructivist outlook of the learning process can be encouraged, in which students learn in an active way in situational, problem-oriented contexts. Witmer and Singer (1998) determined a durable relationship between immersion in the VW and learning and stated that intensification in presence will also escalate learning and attainment.

A further study that evidences the link between immersion in VR and learning gain was that conducted by Liaw et al. (2019), who also found a greater level of student satisfaction. Their sample consisted of 104 university undergraduates (39 females). Significantly greater scores were gained on 11 of the 13 variables investigated using the immersive VR version of the simulation, with the principal variances occurring with respect to presence and motivation. They revealed an affective path in which immersion predicted presence and positive emotions, and a cognitive path in which immersion nurtured a positive cognitive value of the task in line with the control value theory of success emotions. However, as was discussed in Chapter 2, one limitation of this study was that the outcome variables only included self-reporting procedures rather than objective measures. Another limitation was that the sample size was relatively small in respect of the number of aspects that were used in the structural equation model used to test the VR intervention.

In addition to potentially improving student learning outcomes, VR immersion has been linked to learner enjoyment and satisfaction (Csikszentmihalyi, 2000, Tüzün and Özdinç, 2016, Witmer and Singer, 1998, Nunez and Blake, 2001). Several researchers have concluded that a participant with extremely immersive tendencies will feel more present in the Virtual Environment (VE) and relish the experience more than a participant who does not usually become immersed in activities (Nunez and Blake, 2001; Witmer and Singer, 1998). Warburton (2009) highlighted the importance of immersion, stating that the immersive nature of the VW, crossing physical, social and cultural dimensions, can provide a convincing educational experience, predominantly in relation to simulation and role-playing activities. He concluded that immersion in a 3D environment where the augmented sense of presence, via virtual embodiment in the form of an avatar and extensive styles of communication, can affect the affective, empathic and motivational features of the experience.

Again, returning to “interest theory”, Dewey (1913, cited in Slater and Wilbur 1997), argues that students learn via practical experience in organic circumstances and tasks by actively interacting with their setting. This realist experience, argued Slater and Wilbur (1997), could be fostered in virtual reality environments (VREs), promoting higher immersion and enhanced student learning. Furthermore, Milk (2015), emphasised that the finding that amplified immersion can lead to affirmative instructive outcomes is precisely pertinent for immersive VR because the sense of presence experienced by the student can have a potent emotional bearing. Milk (2015) continued by adding that he believes the true power of VR can connect humans to other humans in a profound way that no other form of media can do; he even asserted that it can change people's perception of each other through its immersive properties.

Overall, the precise relationship between immersion and learning outcomes has not been made clear. Whilst Lee et al. (2010) found a direct relationship between immersion and learning outcomes, (when VR is used), Liaw et al. (2019) found that presence plays a mediating role in the relationship. Use of PLS-SEM will enable closer examination of this complex association. On the one hand, Singer et al. (1997) found a statistically significant relationship concerning presence in a VW and spatial learning, while Bailey and Witmer (1994) carried out their study using different measurement tools and found no correlation. Whilst the nature of the precise action of immersion on the outcomes of the VR learning process has not been confirmed, immersion was found to be mentioned highly frequently when researchers evaluated VR learning. The data from the papers reviewed seem to indicate that engagement leads to immersion which in turn leads to knowledge gain (though whether immersion fully or partially mediates the pathway is yet to be proven).

3.6 Conceptual framework proposed

A conceptual framework was proposed based upon the findings of the CR literature review. Based on theories (discussed above) about how mechanisms (in the dominion of the physical (Bonell et al., 2016) interact with context to produce outcomes (in the dominion of the actual), the hypothesis was that in a statistical investigation of outcome gauges (in the dominion of the empirical), students randomly selected to learn about complex diabetic concepts using desktop VR will demonstrate higher short-term learning than students randomly selected to be controls.

It was further hypothesised that in statistical mediation analyses, confidence would mediate the association found between experience and knowledge. There was an indication from the literature (Liaw et al., 2019, Hill, 2017) that prior experience might increase students' knowledge and that improved confidence might then result in heightened student learning outcomes (Schunk, 1995). However, the previous research to base this hypothesis on is inconclusive thus far. The outcome of this hypothesis might shed further light on this questionable aspect.

The second mediation hypothesis drawn from this CR review was that immersion would mediate an association between student engagement and knowledge. This hypothesis has more conclusive evidence from previous research, including evidence from Klevjer (2008), Huang et al. (2013) and Tcha-Tokey et al. (2018). These mediation hypotheses were incorporated into the conceptual framework for testing via PLS-SEM analysis.

The main question that this CR literature review (namely: *"What factors influence student learning when using virtual reality?")* aimed to identify contextual indicators to support explanation of the observed phenomena when students learn using VR. By searching the literature through a CR lens, the following factors were deemed to influence student learning: experience, immersion, engagement, confidence and knowledge. Having considered the literature surrounding influential factors, Table 3-4 illustrates two main

CPMAO configurations that evolved from the above Critical Realist Review of the literature that can be tested via a CR evaluated RCT and PLS-SEM analysis.

Table 3-4 Proposed CPMAO Configurations to be Tested

	Contextual Mechanisms (CM)	**Programme Mechanisms (PM)**	**Agents (A)**	**Outcome (O)**
Explanation/definition of terms **Drawing upon Porter (2015) and Archer et al. (2013)**	The established mechanisms faced by programme initiators.	The range of mechanisms contained in programmes designed to countervail against the powers of contextual mechanisms	The ability to think, reason, imagine, plan and believe, which leads to acting. Participant interpretations and responses.	Outcome produced by change in approach.
CPMAO configurations				
CPMAO#1	Amount of diabetic nursing experience and relevant computing EXPERIENCE required to complete exercise and understand complex diabetic concepts.	Affordances of VR including opportunities to practise via experiential learning, and affordances of computing exercise, e.g., immediate feedback given and the chance to repeat exercise	CONFIDENCE boosting via feedback	Improved KNOWLEDGE of complex diabetic issues
CPMAO#2	Participant ENGAGEMENT level with the learning content	Affordances of VR, e.g., visualisation of abstract concepts	Participant IMMERSION with the desktop VR	ENJOYMENT of the learning activity

The review identified some commonalities, chiefly in relation to the mechanisms at play. This was expected, given that mechanisms in this context are principally identified with human agency (Porter, 2015), and are, consequently, anticipated to be shared by many of the learners involved. As is common with other realist reviews (McConnell and Porter, 2017), contextual mechanisms were addressed in a smaller amount of depth in the literature. However, some insight was gleaned surrounding mechanisms that appeared to promote the successful implementation of VR in educational settings. Variables (or mechanisms) were identified in terms of the variation in learners' levels of confidence, engagement, immersion and experience when learning in VR, which contributed to their subsequent knowledge gain. These are all very much intrinsic to the individual learner and therefore address the "to WHOM" is VR learning beneficial" element of this research very well.

This review resulted in a few key findings that, while important, are not unanticipated given the recognition of VR within education. Nevertheless, the findings do highlight the significance of detecting the synergistic consequence one aspect of the human experience has on others and the ability of VR to tap into this interaction, thereby providing an exciting way to ignite all aspects of learning. For example, despite the many advances in VR technology research has shown that not all learners benefit from the affordances that it provides (Mayer et al., 2008, Cobbett and Snelgrove-Clarke, 2016).

The CR review has highlighted that VR learning is highly complicated in nature as it incorporates many overlapping and related mechanisms (or latent variables), on which researchers have not agreed how, or in what ways, they work to influence student learning; and this provides sound reasoning to pursue a PLS-SEM approach to the analysis of the effects. Furthermore, despite the prospect that different people will benefit from VR (and its effects, e.g., immersion) in different amounts (or not at all), there is sparse information in the literature about the types of individuals who would most profit and few specific interpretations of how VR has helped or not helped to improve the experience of learning and in what particular circumstances.

There remains a significant gap in the research knowledge surrounding the nuances of the impact of VR learning at an individual level. Despite the scarcity of evidence in relation to VR learning and interpretations of students and academics who experience it, this chapter has endeavoured to identify the underlying mechanisms and facilitating contexts that can be tested in the empirical element of the current research in order to provide a more full-bodied understanding of the active ingredients involved. Contextual mechanisms for facilitating the implementation of VR did not overtly emerge from this CR review. However, open-ended questions in the survey and FGDs and interviews will provide participants with the opportunity to raise any key contextual mechanisms that surface as they experience VR learning.

3.8 Limitations of the Critical Realist Literature Review

The strength of the findings in this review is dependent on the strengths of the outcomes in the studies that were reviewed, along with the comprehensiveness of their information about hypothesised programme mechanisms, contextual mechanisms, and human agency. While contextual mechanisms were not identified in depth in this review, it is acknowledged that such data might be accessible in the grey literature that was not explored and as implicit knowledge within the thoughts of experienced VR educators. The underlying mechanisms highlighted in this review deliver a basis for a deeper understanding of what works, for whom (in relation to learners' varying immersive tendencies, engagement levels, confidence and prior computing and nursing experience), and to a very limited degree, in what circumstances.

As in all realist reviews, judgements on all inferences had to be made from the included literature. However, all steps in the appraisal process and analysis were presented as transparently as possible, so that others can clearly see how assumptions and theories of VR learning in HE were arrived at.

It is recognised that restricting the search to English language papers may have resulted in some applicable articles being unexploited. However, in line with the principles of the realist review, inclusive purposive searching was conducted to arrive at a “maximum variety sample” that could sufficiently test the theories (Jagosh et al., 2014). Therefore, inferences made from these studies may not be as robust as potentially they could have been, hence transferability of the intervention theories to other countries is limited.

3.7 Chapter Summary

This critical realist literature review offers discernments into the variety of factors considered to add to an enhanced learning from a VR intervention. A critical realist literature review involves a specific methodology − aiming to build theory concerning the mechanisms at work in different organisational contexts, in this case educational settings and HE. Consequently, this chapter has attempted to gaze beneath the surface of VR learning interventions and begin to realise how they act, conveying the importance of how processes encourage and combine with the agency of the people involved.

By way of summary, the CR approach to illuminating the efficiency of an intervention involves revealing outcomes in terms of a permutation of contextual mechanisms, programme mechanisms and agency. In line with this approach, the CR review reported in this chapter acknowledged the alleged mechanisms embedded in VR that have been identified in the literature as contributing to learning. It examines these mechanisms in terms of their anticipated effects upon human agency and the impact of the implementation of VR interventions in HE settings. The four key themes emerging from the CR review, and any information about the relationships between them, shaped the conceptual framework of interrelated factors that, it is hypothesised, contribute to knowledge when learning with VR.

Chapters 2 and 3 reported two distinct literature reviews. The findings of both have been distilled here in order to clarify the line of argument within this thesis. Firstly, direction for the methods selected, described and justified, was provided by the conclusion of the systematic review of the literature (Chapter 2) on VR technology use in HE. This conclusion argued that whilst measurements have become more objective, stronger evaluation (e.g., larger scale, and using controls and pre-testing) is necessary to confirm the efficacy of VR use in HE. Secondly, the CR review highlighted that VR learning is highly complicated in nature and it provided sound reasoning to pursue a PLS-SEM evaluated RCT approach to the analysis of the effects. The CR literature review identified latent variables for the conceptual model (comprised of two main CPMAO configurations) that was evaluated via a mixed methods approach (described in Chapter 4).

Chapter 4 Chapter Four Research Methodology and Methods

4.1 Synopsis

This chapter details and provides justification (considering the Systematic and CR literature reviews) for the methodological approaches taken in order to address the main aim and objectives of the thesis. In this chapter the overview of the study is discussed, along with the choice of analytical tools, and the quantitative-qualitative methods selected, including how these methods were combined. The hypotheses generated to test the RQs are also presented.

Much of the focus of the chapter is aimed at rationalising and validating the pairing of CR with PLS-SEM as a research methodology for evaluating the efficacy of VR nurse education. The process used for statistical data analysis (namely PLS-SEM) will be described at length because the pairing of PLS-SEM with a CR approach to RCT evaluation represents the main contribution to knowledge claimed in this thesis.

4.2 Methodology

The paradigm adopted for this thesis is a CR approach (Cantor, 1982) which has been used to conceptualise how latent variables (or mechanisms) influence VR learning in HE. CR assumes that outcomes are the result of a combination of mechanisms that have the power to effect change: those contained in the intervention; those embedded in the contexts in which the interventions are implemented; and those that are entailed by individual preference and decision-making. Outcomes will be contingent upon the permutation of mechanisms involved in any given situation.

Porter (2015) advocated a methodological dissimilarity between approaches formulated to recognise the mechanisms embedded in an intervention and its social context, and those intended to expose the experiences, interpretations, and reactions of the actors involved. For the purposes of this current study, an RCT was used as a strategy to identify and test mechanisms entrenched in an intervention (survey linked) and the experiences, interpretations and responses of participants was sought via FGDs and an interview.

A brief background to CR philosophy was provided in Chapter Three; however, CR is presented here in line with ontological and epistemological considerations. According to Wynn and Williams (2012), CR has developed as an alternative philosophical paradigm for carrying out research, particularly within the social sciences. Bhaskar (1989) and others have positioned CR as an alternative to positivist and interpretivist paradigms. CR methodologies can be useful for holistically exploring very complex systems' phenomena and for developing more in-depth causal explanations taking into consideration a broad range of generative factors.

CR is an ontologically based philosophy of science that aims to answer the question "What must reality be like to make science possible?" (Danermark, 2002). The main application of CR is on the supposition that the concepts shaped by scientific research must focus on the objective reality encompassing the world, despite our lack of understanding about this reality and related human fallibility of knowledge. For this reason, Cruickshank (2004), described CR as "ontologically bold", but epistemologically cautious. Hence, CR research concentrates on responding to the question of what reality must be like to illuminate the expressions of given sets of events.

From an ontological point of view, CR accepts that the world is real and continues to exist regardless of whether humans are aware of it (Bhaskar, 1989). Easton (2010) asserted that critical realists consider the world to be socially constructed but that real-world mechanisms also influence social phenomena. Hence, Collier et al. (2019) synthesised the fact that the explanation of social phenomena is made up of both subjective interpretations (interpretivist) and causal mechanisms (positivism). In line with positivists, critical realists state that there are regularities (and indeed demi-regularities) and real-world causal mechanisms that influence such social phenomena. However, since the world can be an "open system" the influences of mechanisms on social phenomena are context-reliant (Shaffer, 2015). The intervention evaluated by this thesis was implemented into an "open system", as a large-scale lecture theatre can certainly be such a system. The ontological and epistemological comparisons of research paradigms are presented in Appendix Nine, which evidences CR as a centralised approach.

CR accepts a stratified ontology comprised of three levels (Mingers, 2004). The chief aim of CR (Fletcher, 2017) is:

> *"to explain social events through reference to these causal mechanisms and the effects they can have throughout the three layers of reality".*

The domain of the real comprises the generative mechanisms, which according to (Eacott, 2015) are: "causal structures that generate observable events". Such mechanisms have an "intransitive" objective reality that is independent of humans' thoughts or values they operate beyond social constructs. Appendix Ten shows the three stratified levels.

This thesis agrees with Bhaskar's (1989) belief that researchers conduct observations in the empirical domain, yet the events take place in the actual domain and might not be completely observable by researchers. Furthermore, there are usually several causal powers involved and therefore it is the combinations of these powers that influence event outcomes. Subsequently, events can be perceived to be signs to comprehend these powers in the domain of the real, though such powers can be difficult to separate from the effects of the context (Houston and Montgomery, 2017). These overlapping and nuanced relationships lend themselves well to analysis via PLS-SEM (as will be fully discussed later within this thesis) as PLS-SEM is fundamentally about the relationships between and combinations of causal powers and their contexts.

Epistemological conventions focus on the idea of what amounts to satisfactory truth by stipulating the criteria and process of evaluating truth claims (Chau Kwong et al., 1998). Then conventions regulate the presentation of knowledge claims, the evaluation of the truth or validity of claims, and how such assertions will be assessed in relation to current knowledge. CR aims to interpret the objective reality through experiences observed by participants and data analysis. Therefore, CR knowledge claims focus on those constituents of reality (e.g., mechanisms and structures) which must be present for experiences/events to have occurred. Such knowledge claims are drawn from numerous epistemological expectations intrinsic to CR, including socially mediated knowledge, explanation of understanding, and the unobservability of multiple possible mechanisms (Bhaskar, 1989, Collier et al., 2019, Easton, 2010).

Within a research project, the methodological strategies chosen to understand the world must align with both the ontological and epistemological underpinnings (Trochim et al., 2008). These three aspects form the theoretical paradigm of this thesis and guide the way this research can be understood. CR lends itself to a variety of research methods, including both qualitative and quantitative (Danermark, 2002). The choice of methodology depends on the ability and complementary nature of varying methods to communicate differing types of knowledge about generative mechanisms (Blom and MorÉN, 2011). Due to a CR focus on explanation, it is argued in this thesis, that the principle of retroduction (from events to mechanisms), becomes relevant.

Rather than an inductive or deductive approach, CR aligns with abduction and retroduction (Bhaskar, 1989, Fletcher, 2017). Some authors have identified differences between these two approaches, e.g., Fletcher (2017), who argues that process of abduction – also known as theoretical redescription – involves redescribing empirical data using theoretical concepts. Others, including Mingers (2004), treat abduction and retroduction as indistinguishable:

> *"retroduction (this is the same as "abduction" as developed by Peirce in contrast to induction and deduction) where we take some unexplained phenomenon and propose hypothetical mechanisms that, if they existed, would generate or cause that which is to be explained."*

Both abduction and retroduction are viewed as appropriate methods to develop hypothesised causal explanations grounded in unobservable generative mechanisms (Blaikie and Priest, 2017). In this thesis the term retroduction is felt to be the most suitable, as it is concerned with not just identifying relationships of cause and effect, but also establishing what underpins these and what creates invariability among phenomena, working back from findings and conclusions to establish the propositions and assumptions on which they are based.

Retroduction permits researchers to move between the knowledge of empirical phenomena to the creation of explanations (or hypothesising) in ways that hold "ontological depth" and can provide some indications on the existence of unobservable entities (Downward and Mearman, 2007). This makes it conceivable to understand how things would have been different, e.g., if those mechanisms did not interrelate in the way they did. Within retroduction, “we take some unexplained phenomenon and propose hypothetical mechanisms that, if they existed, would generate or cause that which is to be explained,” (Mingers, 2004). It can be a creative process whereby the researcher proposes various explanations which describe causal mechanisms, founded within social structures that must be in existence for observed events to be produced. The goal of retroduction is to detect the most comprehensive and rationally convincing justification of the observed events given the context and its conditions; this was also a main goal of the research reported in this thesis.

4.3 Justification for the Methods and Approaches

Justification will now be provided for the selection of the following methods and approaches implemented and described in this thesis, including the use of mixed methods, the CR RCT, the pairing of PLS-SEM with an RCT, and the use of VR technologies to innovate nurse education.

4.3.1 Justification for using Mixed Methods

Mixed methods research was selected to conduct this research. There were several reasons behind this choice. Firstly, Collins et al. (2017) provided a wide-ranging list of motives for conducting mixed methods research; each of these resolves was grouped under four justifications: contributor enrichment, instrument reliability, treatment veracity, and significance enrichment. Secondly, adding quantitative aspects to otherwise qualitative research should lead to more objective and robust learning outcome measures, whilst qualitative research adds colour and depth to numerical data. Finally, as Justice et al. (2007) noted, innovations in HE tend largely to be under-evaluated.

Not only were mixed methods chosen, but a fully integrated form was selected. Cohen et al. (2017) contended that the stronger the combination of methods and their integration at each phase, the better the outcomes of the mixed method approach. Hence, the intention here was to integrate the research at all possible stages, including paradigm, methodology, RQs, instruments, sampling, data analysis and interpretation, and reporting of outcomes and discussion (Creswell, 2018). It is argued that the use of VR to support learning in HE acts in a complex manner and involves several interacting mechanisms, this is supported by the surprisingly slow uptake of embedding this innovative technology within curricula (Wimpenny et al., 2012). Due to the complex nature of the research, mixed methods were deemed to be most suitable.

Whether a CR approach complements the use of mixed methods was also considered. CR does not commit to a single type of research but rather endorses a variety of quantitative and qualitative research methods (Zachariadis et al., 2013). This critical methodological pluralism is not taken flippantly but has its foundations strongly in the ontological and epistemological assumptions of CR, thus conserving a durable connection between meta-theory and method (Danermark, 2002). It is felt that an RCT, from a CR perspective, will effectively blend both qualitative and quantitative elements and assist in uncovering the more nuanced views and experiences of those involved in using the VR technology. It is these views and experiences that will pinpoint future direction for the implementation of such HE innovations.

4.3.2 Justification for using Critical Realism

This thesis proposes CR as a more integrated approach to holistic evaluation of complex interventions either in health care or, as in this case, education.

CR is particularly apposite for mixed methods research, permitting it to have veracity and coherence as it addresses multiple RQs by using a variety of research methods, counter-attacking the critique of methodological incommensurability (Walsh and Evans, 2014). Blackwood et al. (2010) discussed the "philosophical tension" involved in merging RCTs with qualitative approaches. They go on to assert that CR provides the promise of combining these two experiential strategies logically, and illustrate the point in this comment:

> *"The RCT can be used to ascertain whether, all other things being equal, a particular causal mechanism (intervention) is efficacious, while realistic evaluation can establish what effect the interaction of other mechanisms operating in the open contexts studied has upon its effectiveness, and identify which mechanisms promote, and which inhibit that effectiveness."*

Using a CR approach, hermeneutical appraisal of interventions requires focusing on both the social mechanisms they involve and the reactions to these by individuals affected by them. Hence, the purpose of the CR evaluation of an RCT is, in part, to establish the degree to which social agents experience interventions as encouraging or constraining learning. An RCT methodology by itself can fall short of examining the bigger picture. For this reason, amongst others, numerous researchers have argued that the RCT, with its aptitude to recognise the efficacy of an intervention within the restrictions of a closed system, is an essential but not adequate methodology, and needs to be shared with realist-based examinations of individual experience and social context (Bonell et al., 2016, Porter, 2015).

4.3.3 Justification for a Critical Realist Randomised Controlled Trial Approach

The justification for the suitability of a CR RCT approach, based upon the systematic literature review findings, will now be deliberated. As part of the systematic review, key terms and main affordances/challenges of the technologies under discussion were identified. Exploration of the ways in which VR technologies used in HE have been evaluated highlighted the fact that the most rigorous studies combined reliability and validity approaches based upon Ader et al. (2008), Creswell (2018) and Cohen et al. (2017), notably: pre-testing, use of a control and appropriate sample sizes. The systematic review (Chapter Two) argued that whilst measurements have become more objective, stronger evaluation is necessary to confirm the efficacy of VR/VW/AR use in HE. Therefore, for this study, the aim was for a strong study design, including: randomisation, control, pre-testing, objective data collection, appropriate sample size (over 55 in each group according to Hair et al., 2018), a heterogeneous sample population (BU and Yeovil student cohorts), balanced reporting of findings (including negative findings and limitations), and statistically significant findings.

Turning now to justification for the use of an RCT to evaluate the VR intervention, Ripoll et al. (2018) stated that traditionally, VR learning has been assessed by examining the self-reporting of user experiences. Where quantitative measures have been used, including RCTs, there have been few that have had large sample sizes or heterogeneous populations (Allcoat and von Mühlenen, 2018). RCTs are the gold standard for providing evidence on the effectiveness of interventions (Sibbald and Roland, 1997). Lohre et al. (2020) recommended that RCTs be used to evidence the transfer of clinical skills when using VR. The advantages of RCTs include: that randomisation minimises the influence of both known and unknown prognostic variables on treatment outcome, RCTs can demonstrate causality, and quality control of intervention and outcome assessment, and finally that RCTs provide the strongest empirical evidence of treatment efficacy (Sibbald and Roland, 1997).

However, limitations of RCTs include that the required study power might not be met, and that generalisability may be low. There has also been criticism concerning the limitations of RCTs for use in evaluating the success of interventions conducted in open systems (such as a lecture theatre setting) (Corry et al., 2019). For these reasons, whilst an RCT approach has been selected for the current study, rigour has been optimised by striving for an appropriate sample size and using heterogeneous populations. Furthermore, a CR evaluation approach will be combined to address the limitations of conducting such research in an open system. Recent advances in CR approaches towards RCT methods have facilitated more holistic investigations of open system interventions (Porter et al., 2017). In practice, this means that the research will not simply try to find out "what works?" or "does it work?" but additionally, "what works, for whom, where, when and in what circumstances?"

Qualitative work in conjunction with RCTs, remains rare, poorly assimilated, and often has major methodological inadequacies (Ferguson et al. 2016). Therefore; there is a need to fully integrate qualitative approaches when combining them with evaluations of RCTs. Juden (2014) asserted that realist RCTs ought to:

> *1. cultivate a clear causal model of the outcomes of importance within the test population to enable: a. an overt explanation of intervention causation b. testing of that explanation*
>
> *2. measure pathway variables to inform causal theory*
>
> *3. appraise the diverse constituents of multifaceted interventions independently as well as in combination*
>
> *4. add to the development and iterative reassessment of CMO typologies*
>
> *5. be combined into a mixed-methods approach to data generation*
>
> *(and)*
>
> *6. openly outline the normative structure from which the intervention pulls its reasoning.*

A qualitative angle could enable evaluation of various interacting mechanisms. Several researchers have highlighted that various mechanisms might interact when learning via VR. For example, according to Liaw et al. (2019) there is incomplete empirical evidence of the affective value of VR, and a reduced amount of research that examines the psychological mechanisms by which additional immersion impacts students' motivation, or whether it could accelerate self-regulation and performance in the learning process. In a similar vein, Villalta et al. (2011) asserted that in order to entirely comprehend how students process a VR environment, it is necessary to consider their emotional responses to that environment .This is particularly significant when crafting educational material, such as VR supported learning activities, as an understanding of the principal mechanisms that influence learners' perceptions and motivations can direct the optimal development of VR simulations, which integrate emotional deliberations that can lead to amplified use and improved outcomes (i.e. learning gain) (Liaw et al., 2019).

Researchers have become increasingly critical of RCTs and have started to reappraise how RCTs can be effectively evaluated. This thesis has applied realist principles to RCT methodology. Two main approaches have been adopted, namely realist evaluation (Pawson and Tilley, 1997) and critical realism (Bhaskar, 1989, Archer et al., 2013, Bazzaza et al., 2016, Bonell et al., 2016). CR has been criticised on the grounds that the conception of mechanisms as 'a function of the interactions between intervention resources and responses of participants' (Marchal et al., 2012) is contrary to the sort of correlational analysis that is intrinsic to RCT methodology. One aim of this thesis is to support the CR approach (Blackwood et al., 2010; Porter et al., 2017) by indicating how the application of PLS-SEM has the prospect to overcome the challenges of applying statistical analysis to the complex configurational approach to causation advocated by CR.

A further limitation of RCTs is that, in contrast to classic experiments, they are applied in semi-open systems that include agents capable of interpretation and choice. However, Marchal et al. (2012) asserted that RCTs can be regarded as epidemiological proxies that replace probabilistic controls over extraneous factors for experimental closure. Nevertheless, the main outcome of RCTs does not relate to the power of a unique mechanism but to that of a configuration of mechanisms, hence creating epistemological difficulties. In terms of the intervention, it can be tricky to differentiate the causal contribution of each mechanism. In relation to context, it can be a challenge to account for the consequences of contextual variability. Control over participant variability (agency) is a further issue. One possible CR solution to these issues is via two additional research strategies. The first comprises the development and testing of realist hypotheses about the way the intervention works (e.g., its rules), its context, and the relationships involved (Porter et al., 2017). The second encompasses qualitative exploration designed to expose how agents experiencing the intervention's workings interpret and respond to it.

With regard to the philosophical compatibility of RCTs and CR, Hall and Sammons' (2013) definitions of approaches to regression analysis demonstrate striking consonance with CR conceptions of causation. They implied that two or more concepts, "*work together*" or, "*have a combined effect*" in producing a third, demonstrating relevance to the CR notion of configurational causation. In addition to argued compatibility of CR with RCTs, this thesis argues that CR is also compatible with a PLS-SEM analysis of such RCTs. It is proposed in this thesis that, philosophically, the notion of "mechanisms" within CR research, and the notion of "variables" within PLS-SEM research have enough commonalities to be considered as similar enough for the pairing of CR with PLS-SEM to be justifiable. For example, if there is a "mechanism" (Ishaq et al., 2019) it must "act" on something(s). If something(s) is/are acted upon, their properties must change. Those properties, measured or latent, are variables.

4.3.4 Justification for the pairing of PLS-SEM with the CR RCT

Justification for the pairing of PLS-SEM with the CR RCT will now be discussed. It is argued that a CR evaluation on its own might not succeed in holistic evaluation of a VR intervention. For example, Gilmore et al. (2016) highlighted the limitations of most realist evaluations, in that they incline towards the qualitative end of the spectrum, and any quantitative analysis concentrates on outcomes, tending to either be descriptive or utilise hypothesis analysis to measure the statistical significance at all stages of the intervention implementation. A small number use more advanced statistical modelling techniques, such as regression, e.g., Ebenso et al. (2019). These practices could supplement the evaluation of qualitative data, e.g., by helping to elucidate the relative importance of a range of configurations that lead to a similar outcome.

One such statistical modelling method is Structural Equation Modelling (SEM). It is a multivariate investigative approach used to concurrently examine and approximate complex causal relationships amongst variables, even when associations are theoretical, or not directly evident (Williams et al., 2009). Simultaneously uniting factor investigation and linear regression models, SEM permits the investigator to statistically scrutinise the associations amongst theory-based latent variables and their indicator variables by computing directly discernible indicator variables (Hair et al., 2014).

SEM is a well-known quantifiable technique which has both a quantifiable and structural component (Stein et al., 2017). The measurement constituent permits proof of identity of unobserved, or latent, variables. For instance, patient “empowerment” is an unobservable notion, but could be recognised from numerous observed variables, such as confidence in knowing when to seek help (Ford et al., 2018). These types of latent variables (such as patient empowerment) are categorised as reflective measures because patient empowerment leads to confidence in knowing when to seek help.

Formative measures are the opposite and occur when the observed variable causes the latent variable (Bollen and Bauldry, 2011). For example, transport options (formative variable) may be determined by car ownership. Reflective measures are like the realist mechanisms (utilised in realist methods) because they are typically conceptualised as being unobservable (Astbury and Leeuw, 2010). Realist mechanisms could therefore be measured using the concept of reflective measures (Ford et al. 2018).

SEM is a suitable vehicle for combining with RCT evaluation because it enables not only the examination of variables (and relationships between variables), but also identification and exploration of latent variables (e.g., unobserved variables) (Hair et al., 2018). Furthermore, Byrne and Uprichard (2012) argue that causal accounts require the examination of interventions from a systems viewpoint with a case-based (i.e. configurational), not a variable-based positioning. The emphasis is firmly grounded in the ways that variables (or in CR terminology “mechanisms”) interact with one another, rather than considering them in isolation with the notion that they do not affect each other and/or that they act independently. A CR approach towards RCTs (and for that matter, systematic reviews) can also counteract the fear that they do not produce results that are directly pertinent to all patients and all situations. To aid external validity, they should be designed and reported in a way that allows learners and educators to judge to whom they can reasonably be applied (Rothwell, 2006).

With the above assertions in mind, it is argued that the CR element of this research not only supports the RCT angle taken, but moreover, it aligns very well with PLS-SEM, which is, in its simplest form, an evaluation of the strengths and directions of relationships between variables. For example, Bonnell et al. (2016) asserted that a realist evaluation concentrates on evolving, refining and analysing theories concerning how interventions deliver resources which participants use to activate mechanisms that interrelate with context to produce outcomes. This is useful both in emphasising that evaluative studies should concentrate on the analysis and modification of intervention theory (rather than simply certifying interventions as effective or not). It is also supportive in providing a foundation for appreciating the significance of context, and for drawing on experimental data to deliberate how context might affect the application and effects of interventions in new situations.

A further justification for the pairing of PLS-SEM with CR is the concept of generative indicators and the idea of evaluating them configurally. The powers of natural generative mechanisms (the realist descriptor for natural laws) are what they are (Porter, 2015), independently of human activity. The purpose of experiments is to recognise those powers by generating configurations of events by means of closure. This conceptualisation moves the examination of causation past correlational investigation, and configurational usage of indicators (CPMAO) can be applied as opposed to isolated variable-based examination (Porter et al., 2017).

According to Ford et al. (2018), the structural component of SEM, quantifies the association between latent or observable variables along a pre-specified route using regression procedures. Although CMOs (Context, Mechanism, Outcome) are arrangements, not correlations, they do have a regular consecutive order of C-M-O and therefore are theoretically responsive to measurement. As best is as known, Ford et al. (2018) is the first study to explore CMO configurations using SEM. SEM permits each CMO configuration to be measured and equated to assess relative strength. Moreover, it enables measurement of CMO configurations within realist philosophy; supplementing qualitative data and descriptive quantitative techniques in realist appraisals to sustain interpretations about powers of associations (Ford et al. 2018). In this thesis it is proposed that SEM (and specifically PLS-SEM) could therefore be used to compliment Porter's (2015) modified equation approach to CR evaluation, namely Contextual Mechanisms + Programme Mechanisms + Agency = Outcome (CM + PM + A = O).

Before continuing, it is helpful to describe Partial Least Squares-Structural Equation Modelling (PLS-SEM) in full. PLS-SEM is a specific type of SEM. It is a causal-predictive approach to SEM that stresses prediction in estimating statistical models, whose structures are intended to provide causal explanations (Sarstedt and Cheah, 2019). Wold et al. (2001) explained the situations in which PLS-SEM was intended for use:

> *"PLS is primarily intended for causal-predictive analysis in situations of high complexity but low theoretical information."*

PLS-SEM permits critical investigative research (Avkiran, 2018). According to Lowry and Gaskin (2014) PLS-SEM is a procedure well matched to evaluating multifaceted prognostic models. Dissimilar to the covariance-based SEM approaches, which necessitate a multivariate normal distribution of the observed variables, PLS is founded on the resampling measures of bootstrapping (Henseler and Hubona, 2016).

Bootstrapping is a non-parametric resampling procedure that assesses the variability of a statistic by evaluating the variability of the sample data rather than using parametric assumptions to assess the accuracy of the approximations (Streukens and Leroi-Werelds 2016). Each of the desired samples is obtained by sampling with replacement from the original data in a way that every bootstrap sample contains as many cases as there are in the original data. There are several advantages of bootstrapping. It is a widely applicable (the number of potential bootstraps is almost limitless), non-restrictive (in relation to assumptions about the data) and transparent method that requires little knowledge of probability theory.

Bootstrapping allows investigators to model and approximate compound cause-effects association models with both latent (denoted as circles in Figure 4-1) and observed variables (denoted as rectangles).

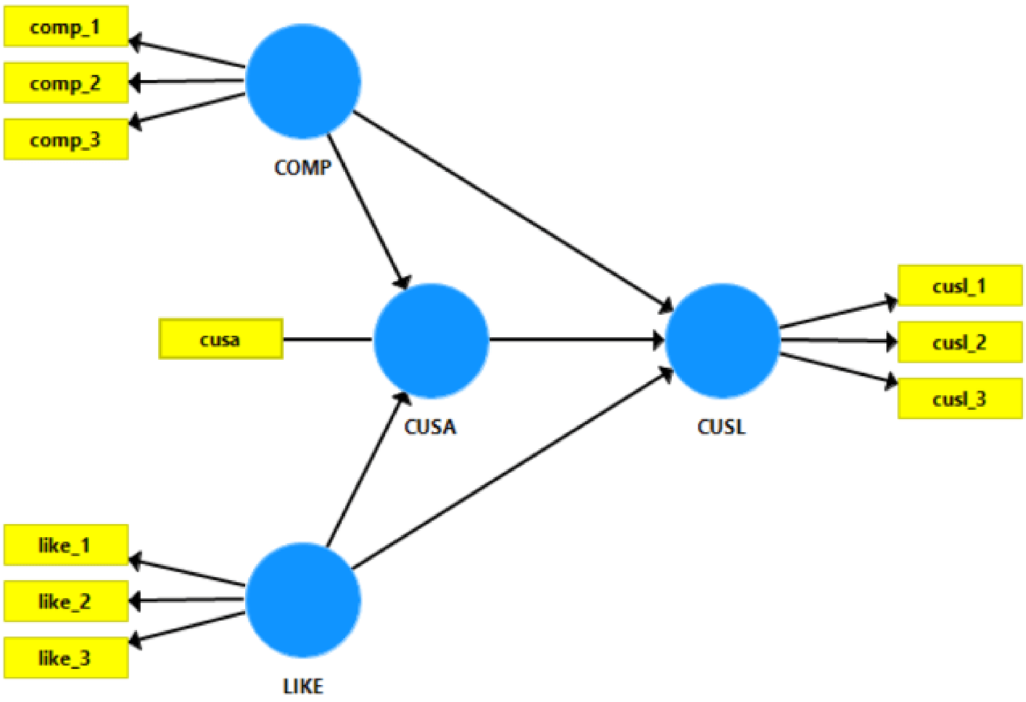


Figure 4-1 Smart PLS diagram showing latent variables (denoted as circles- e.g. perceptions and attitudes) and observed variables (denoted as rectangles- which relate directly to the survey questions). This diagram is the example provided in Hair et al., (2016).

The latent variables exemplify unobserved (i.e. not directly measurable) phenomena such as perceptions, attitudes, and intentions. The observed variables (e.g., responses to a survey) are used to represent the latent variables in a statistical model. PLS-SEM estimates the relationships between the latent variables (i.e. their strengths) and regulates how well the model explains the target constructs under examination (Hair et al., 2017). PLS-SEM is the appropriate approach when the research entails theory development (Hair et al., 2016).

When considering complex research models, the various flexible analysis possibilities, limited suppositions, and user-friendliness of PLS-SEM make it the “holy grail” for advanced methods (Mathews et al., 2018). For example, PLS-SEM can: (1) establish data equivalence via the three stage Measurement Invariance of Composite Models Procedure (MICOM) process (Henseler and Hubona, 2016), to minimise measurement error, (2) detect the significance and performance of precursor constructs to target areas for further research (Hair et al., 2016), and (3) unearth unobserved heterogeneity so that structural and measurement models can be scrutinised either at the individual group level or the aggregate level (Matthews et al., 2018).

An added advantage of the PLS-SEM method is the unlimited integration of latent variables in the path model that either draws on reflective or formative measurements models (Ringle and Sarstedt, 2016). PLS path models comprise three constituents: the structural model, the measurement model and the weighting scheme. Path coefficients are the main outcomes of PLS-SEM which quantify the hypothesised relationships within the structural model (Riou et al., 2016). While structural and measurement models are constituents in all kinds of SEMs with latent constructs, the weighting system is specific to the PLS method.

Drawing upon the findings of the CR literature review reported in Chapter 3, two CPMAO configurations were proposed for the CR-RCT for the current study. The two configurations have been set out along with associated data collection and evaluative methods based upon Porter’s (2015) recommendations (Table 4-1).

Table 4-1 CPMAO Configurations with associated data collection and evaluative methods

CPMAO configurations	**Contextual Mechanisms**	**Programme Mechanisms**	**Agents**	**Outcome**
CPMAO1#	Amount of diabetic nursing experience and relevant computing EXPERIENCE Required to complete exercise and understand complex diabetic concepts	For example: Affordances of VR including opportunities to practise via experiential learning, and affordances of computing exercise, e.g., immediate feedback given and the chance to repeat the exercise	CONFIDENCE boosting via feedback	Improved KNOWLEDGE of complex diabetic issues
CPMAO2#	Participant ENGAGEMENT level with the learning content	Affordances of VR, e.g., visualisation of abstract concepts	Participant IMMERSION with the desktop VR	ENJOYMENT of the learning activity
Recommendation of how to evaluate these (Porter, 2015).	Testing of hypotheses about the mechanisms embedded in the existing social context (CM)	Testing of hypotheses to identify the mechanisms embedded in the intervention designed to countervail against what are identified as problem mechanisms in the social context	An examination of how agents interpret and respond to these mechanisms	Recording and explanation of changes in rates And study of the consequences of interventions for the lives of those affected by them
Data Collection via:	Surveys and FGDs	Surveys and FGDs	Surveys and FGDs	Surveys and FGDs
Approach	QUANT	QUANT	QUAL	QUANT/QUAL
Method used to evaluate each strand.	PLS-SEM of RCT	PLS-SEM of RCT	PLS-SEM of RCT Open-ended survey questions FGDs Interview	PLS-SEM of RCT Open-ended survey questions FGDs Interview

4.3.5 Justification for using Virtual Reality to Innovate Nurse Education

Finally, justification for the choice of a VR innovation for Nurse Education will now be detailed. Having reviewed the relevant literature surrounding the use of VR in HE; and from experience gained from assisting with TEL sessions for the general public, physiotherapy students, midwifery students and college lecturers using a 360 degree camera, AR and VR, the decision was made to use VR to innovate student nurse lectures at BU.

Due to the fact that nursing students would not be on a ward (real or simulated) AR was not deemed to be the most appropriate technology to enhance their learning, because it requires layering computerised information over the top of the physical environment (Kardong-Edgren et al., 2019). The most significant difference between VR and AR is that users of AR are in the physical environment to which the AR relates, whereas users of VR can be located anywhere, including traditional lecture theatres.

Initially, the decision was made to access the VR simulated exercise via students' mobiles phones and using a VR headset such as Google Cardboard. This was because, although BU owns a few Oculus Rift VR devices and increasing numbers of Oculus Go, Nurse Lecturers needed students to either receive training at the same time in large lecture theatres or access the software at home. Moreover, according to Amin et al. (2016) their results indicated that notwithstanding its simplicity and small screen size, Google Cardboard is adept in providing a satisfactory level of immersion compared to Oculus Rift's larger screen size. Accessing VR via a PC or tablet is considered to provide lower levels of immersion (Miller, 2014), for example, there are usually numerous signals indicating the presence of device(s) in the physical world (e.g., use of a joystick or mouse to control the

VE) and that they only accommodate one sensory modality (e.g., auditory, visual, motor/proprioceptive); and stimuli are not spatially oriented.

Having selected low-cost Google Cardboard VR headsets as a comparator for traditional lecture style teaching methods, a survey was sent out to all Second Year Adult Nursing and Mental Health Nursing Students (both of which receive the same Long Term Conditions Unit), to capture data to find out how many of those students own a mobile device with the capability to run the VR software (e.g., iPhone 7 or above, or Android equivalent). The number of students who stated that they owned the phone was 45; this was too low a number to conduct a randomised control comparing VR headset versus traditional methods. Thus, the decision was taken to hold a bespoke group with the following year's student nursing cohort, inviting only those who have a mobile device of the specifications required to run the software, and Immersive VR data were collected at this later date (though was reported outside the realms of this thesis).

These "Immersive VR" data generated from students using the mobile phone application version of the software along with Google Cardboard, would be compared to the control and desktop VR data already collected (though it is acknowledged and reported that this new group would be randomised by mobile phone ownership rather than by student ID number, as applied to the previous two groups – Control and Experimental). Hence, students in the Experimental group accessed the VR training exercise via desktop VR, otherwise known as non-immersive VR. Desktop was found to enhance learning, despite being less immersive than headset accessed VR in a study conducted by Dubovi et al. (2017). However, evidence retrieved by the literature review surrounding the comparison of desktop VR compared to immersive VR (when considering the impact of learning) was otherwise sparse (Srivastava et al., 2019).

4.4 Aims Relating to the Empirical Research

Previous VR research in HE has been inconclusive regarding the impact of previous computing experience (e.g., Childs, 2010), engagement, immersion, and knowledge gain (Makransky et al., 2019). This research intended to address these inconsistencies and provide direction for educators in HE settings. The main aim of the thesis was to assess the effectiveness of CR paired with PLS-SEM as a method to evaluate the impact of VR simulations on undergraduate nurse education. The aim was achieved through two objectives, which were:

1. To determine the effect of pairing CR with PLS-SEM as an evaluative method

2. To determine how using this novel evaluative method can inform our understanding of the impact and future use of VR simulations for undergraduate nurse education

These two objectives are closely related, and the first objective can be also be understood through objective two. The first of these objectives involved pairing CR with PLS-SEM through the application of a VR simulation for nurse education. The findings resulting from this method could then be compared to findings from other studies which had used more traditional methods (e.g., RCT analysed via t-testing). The second of these objectives involved operationalising the research gaps found in the systematic review by developing a VR teaching and learning exercise to be piloted then deployed at BU with a prearranged group of BU students (adhering to BU ethics protocol). The reason for this innovation was to enhance the traditional course curricula. The second objective also involved conducting process and outcome evaluations of the intervention exercise using qualitative and quantitative measures: firstly, by delineating the impact of the teaching and learning exercise for the students and academics involved, and secondly to provide direction for both the scalability and sustainability of similar VR learning systems at BU and in other contexts.

4.5 Hypotheses

Based on theories (discussed in Chapter 3) about how mechanisms (in the dominion of the physical, Bonell et al. 2016) interact with context to produce outcomes (in the dominion of the actual), the hypothesis was that in statistical investigation of outcome gauges (in the dominion of the empirical), students randomly selected to learn about complex diabetic concepts using desktop VR would demonstrate higher short-term learning than students randomly selected to be controls. It was further hypothesised (see Section 3.9) that in statistical mediation analyses, confidence would mediate the association found between experience and knowledge. Moreover, it was hypothesised that immersion would mediate an association between student engagement and knowledge. These investigations were aimed at helping researchers and educators gain a more vibrant and nuanced appreciation of how and why enhanced learning is caused by mechanisms that interact when students use desktop VR, which are triggered and expressed differently in varied, HE contexts.

The hypotheses are set out as follows:

H1a: experience will have a positive influence on knowledge

H1b: confidence will mediate the experience to knowledge relationship

H2a: engagement will have a positive influence on knowledge

H2b: immersion will mediate the engagement to knowledge relationship

H3: the experimental group will be more knowledgeable than the control group

H4: the experimental group will be more confident than the control group

H5: the experimental group will be more engaged in their learning than the control group

From these overarching hypotheses the following specific hypotheses were developed and tested using PLS-SEM analysis (Table 4-2).

Table 4-2 Specific Hypotheses tested as one way of addressing the main thesis aim and objectives

Hypothesis Number	***Specific Hypotheses***
H6	Path CONFIDENCE→KNOWLEDGE, $H_\circ$ β=0; (H_1 β ≠ 0)
H7	Path *ENGAGE→IMMERSE*, $H_\circ$ β=0; (H_1 β ≠ 0)
H8	Path *ENGAGE→KNOWLEDGE*, $H_\circ$ β=0; (H_1 β ≠ 0)
H9	Path *EXP→CONFID*, $H_\circ$ β=0; (H_1 β ≠ 0)
H10	Path *EXP→KNOWLEDGE*, $H_\circ$ β=0; (H_1 β ≠ 0)
H11	*Path IMMERSE→KNOWLEDGE*, $H_\circ$ β=0; (H_1 β ≠ 0)
H12	LV KNOWLEDGE, $H_\circ$ $Gp1_{mean} – Gp2_{mean} = 0$; ($H_1$ $Gp1_{mean} – Gp2_{mean} < 0$)
H13	LV CONFIDENCE, $H_\circ$ $Gp1_{mean} – Gp2_{mean} = 0$; ($H_1$ $Gp1_{mean} – Gp2_{mean} < 0$)
H14	LV ENGAGEMENT, $H_\circ$ $Gp1_{mean} – Gp2_{mean} = 0$; ($H_1$ $Gp1_{mean} – Gp2_{mean} < 0$)
H15	LV IMMERSION, $H_\circ$ $Gp1_{mean} – Gp2_{mean} = 0$; ($H_1$ $Gp1_{mean} – Gp2_{mean} < 0$)

These hypotheses were tested as one way of addressing the main aim and objectives.

4.6 The method

The justification for the methods and approaches used within this thesis have been provided in an earlier section (Section 4.3). This section now provides details of the methods used. As part of the method process, CR evaluation continued in tandem with the trial and entailed three separate though interrelated approaches. The initial phase was to estimate outcomes; the second was to further cultivate and investigate hypotheses about the mechanisms rooted in the intervention and its context; and the third was to expose how the people involved react to the resources and constraints created by the intervention and its context. The latter two strategies involved qualitative methods.

The combination of CR evaluation and an RCT (analysed via PLS-SEM) was a pioneering approach to evaluation methodology and was the first of its kind (as far as is known). It aimed to explain inconsistencies amongst expected and observed results, to comprehend how context impacts on outcomes, and to deliver insights to aid future application. The use of a CR approach was intended to overcome the disadvantages of conducting a trial in real-life open systems (e.g., a university lecture setting). The combination of a traditional RCT approach alongside CR evaluation joins the notion of “is it of benefit/does it work?” with the notion of “what works for whom, how and in what circumstances” (Pawson and Tilly, 1997). The combined process is illustrated in Figure 4-2. By using this CR approach, findings from the current study will provide direction for future scalability of the use of VR technologies in teaching and learning settings at BU and beyond.

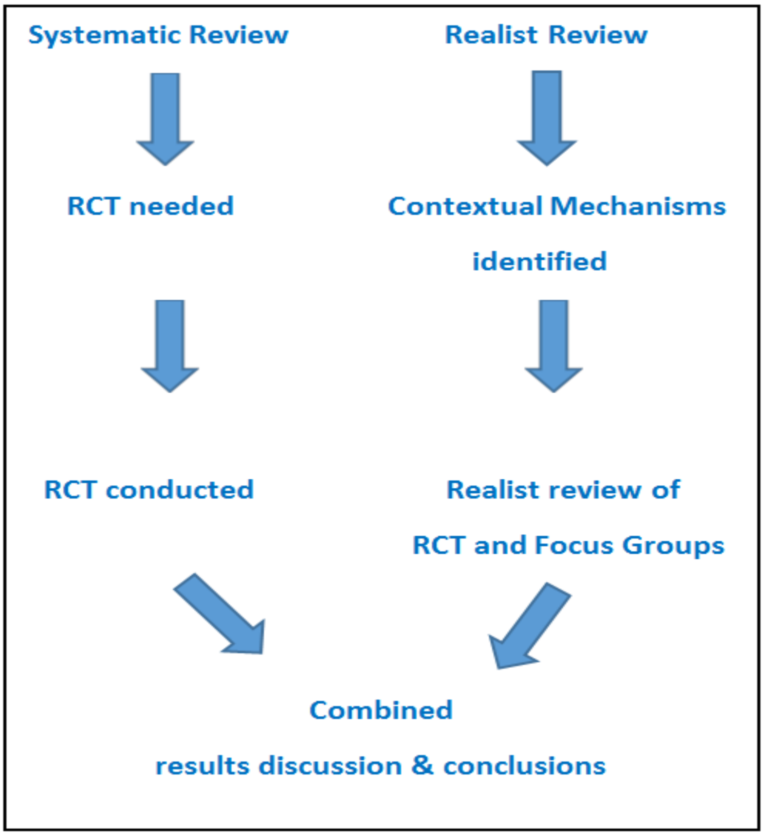


Figure 4-2 The method process followed in this thesis

4.6.1 The characteristics of the sample

All Second Year Adult and Mental Health Nursing students (BU and Yeovil sites) were invited to complete an online consent form and survey. The number of students who agreed to take part in the study was 216 (81% of the total population), of which 171 (64%) completed both the pre and post-test surveys. Participants provided details of their student ID number which was used to randomise students into the Experimental or Control Group.

This research followed Cohen's (1992, cited in Hair et al., 2017) recommendations for sample size calculation, suggesting that when the maximum number of independent variables in the measurement and structural models is five, there needs to be at least 45 participants in each group (in PLS-SEM) to attain a statistical power of 80% for detecting two values of at least 0.25 (with a 5% probability of error). The sample was representative with respect to gender and age.

4.6.2 The Intervention

Participants (BU 2nd year nursing students) were randomly assigned to one of two groups, either the Control Group or the Experimental Group. The Control Group completed a paper version of the hypoglycaemia case study (Appendix Two), which provided redirection if they made an adverse decision in the patient's care. The Experimental Group completed a VR-based exercise, which was developed with the aim of improving student nurse engagement with, and knowledge levels relating to, a diabetes and chronic illness unit of work. The specific focus of the exercise was the recognition of the symptoms and management of hypoglycaemia (low blood sugar levels) in a diabetic patient. The desktop VR version of the exercise (shown in Figures 4-3 and 4-4) was based upon the same case study as the paper version. Knowledge and hands-on experience gained during the first year of the research enabled the arrangement and coordination of the writing up of the deteriorating patient case study, in conjunction with two BU Diabetes Nurse Specialists, ensuring that it could be translated into a feasible VR exercise.

The central education unit, within which the PhD studies were located, allocated funding to support the research. Work was commissioned with Daden Ltd, having benchmarked the sector and ascertained that they could produce a suitable VR exercise. Daden Ltd programmed the VR application based on the deteriorating patient script. After a series of piloting, communication with Daden Ltd and reiteration, the low-cost, proof of concept simulation was completed and was ready for use on students' laptops. The app, or headset, version has also been produced, though is not yet available on all mobile devices. Trials of the headset version compared to the laptop version were planned for the start of the next academic year, the outcomes of which will be published but will not form part of this thesis. Thus, the VR exercise was based on specialist nursing expertise and VR expertise (gained from the author's supervisors and through experience).

For this present study, a design-based research (DBR) approach (Johnson et al., 2017) was used in implementing this lecture theatre-based research. Specifically, the four phases of Reeves' (2005) methodology were intuitively charted: (i) the examination of the problem and literature review; (ii) creation of the simulation solution; (iii) iterative application of that solution into the lecture theatre by appropriate discipline lecturers; and (iv) a loop back to design modification and further iterative analysis to evaluate the development of an effective desktop VR simulation.

Table 4-3 The Design Based Research Approach used in this thesis

Phase One	Phase Two	Phase Three	Phase Four
The examination of the problem	**Creation of the solution**	**Iterative application of that solution**	**Loop back**
Through two literature reviews: -A systematic literature reviews -A Critical Realist literature review Discussion with: Nursing academics, Diabetic Nurse Specialists (DNSs) at the local hospital, BU LTs, a software company	Using knowledge gained from assisting with VR technology workshops at BU. Using expertise from DNSs and academics from BU. Using Diabetes UK guidelines for treating hypoglycaemia. Hypoglycaemia deteriorating patient case study sent to software company to be made into a VR simulation.	Piloting of the software by BU academics and LTs. Application of the VR simulation for Adult and Mental Health Nursing students at BU.	Feedback taken from academics, DNSs and LTs given to the software company for future iterations of the VR simulation.

Figure 4-3 Nurse Avatar's view of the Hypobox and Blood Glucose Monitor

Figure 4-4 Nurse Avatar deciding the next course of treatment

4.6.3 Description of the VR simulation

Students play the part of the nurse avatar who stays within the ward side room as shown in Figure 4-4. The simulation begins with the nurse in charge providing the student nurse avatar with the nursing handover. The handover provides the nursing student with knowledge about the patient’s condition, current medication and observations (e.g., blood pressure, blood glucose, oxygen saturations, temperature, heart rate and respiratory rate) which had been recorded the night before. The handover also detailed the patient’s history, which included the fact that the patient had Type 2 diabetes and had been admitted to the ward with a chest infection and, at admission, hyperglycaemia (high blood glucose). The patient had been given oxygen therapy last night and had received IV fluids, both of which have now been discontinued.

The nurse in charge instructs the student to check the patient’s observation chart and then carry out observations. The equipment in the room (blood pressure machine, blood glucose monitor for instance) makes noises, has time delays and is modelled on nursing equipment from the local hospital. From this point the student takes over the nursing care of the patient seeking advice from the nurse in charge as needed.

When the student approaches the patient, the patient is irritable and very sleepy. The student must make a safe clinical decision about how to react and communicate with the patient. If the student makes unsafe decisions, they are given instant feedback that their decision was unsafe and that they need to think again quickly because the patient is deteriorating. The patient begins to look unwell if the student does not correct the patient's blood glucose quickly using medications from the Hypobox. It is vital that the student realises that although upon admission (the previous night), the patient was suffering from hyperglycaemia (high blood glucose) they are now experiencing hypoglycaemia (low blood glucose). This low blood glucose must be corrected via the administration of Glucotabs before the patient can be given their breakfast or their usual insulin (this is in accordance with the Diabetes UK guidelines 2019 for a patient with this level of consciousness and signs and symptoms). The student must also continue to monitor the patient's blood glucose using the monitor until it has returned to above 4 mmol/L.

When the student measures the patient's blood glucose, they see that it is below 4 mmol/L. They are then presented with clinical decisions (in the form of multiple-choice questions via pop-up text boxes). The clinical decisions include when to seek advice from the nurse in charge, when to retest the blood glucose, what medication to use from the Hypobox, when to give the usual insulin and when to give breakfast for instance. Each clinical decision is followed up with instant feedback so that the student can learn and improve. There is a question at the end that confirms the student's learning about the patient's usual medication and the risk of repeated hypoglycaemia after they have been discharged from hospital. The student can complete the simulation multiple times and the lecturer is sent data analytics about each student's performance.

4.6.4 Piloting of the Software

Piloting of the software was conducted with LTs and academic staff. The process was iterative with any issues fed back to the software company for improvements to be made. This process will be ongoing beyond the life of this thesis, as new cohorts use the software.

Figure 4-5 The Software showing the pop up text boxes used.

4.6.5 Data Collection

As the method included an RCT, a randomisation list was computer-generated, using GraphPad- Randomize. Using student ID numbers, each participant was randomised into either the Control or Experimental Group. If participants were randomised into the Experimental Group but could not bring a laptop to VR experiment day, they were permitted to use a laptop provided by the learning technology support services.

4.6.6 The instruments

On obtaining written informed consent from the nursing students, an online survey was run using Bristol Online Survey (BOS) and students' own handheld devices. Students who did not have their own devices were able to use a friend's device. The survey was anonymous apart from students' ID number (ethically cleared to be used to link students' pre-test to their post-test for analysis purposes).

The design of the surveys (pre and post) was based primarily upon the PLS-SEM conceptual framework which evolved as a result of the CR review of the literature (Chapter Three). Questions were also drawn from tried and tested questionnaires, in order to optimise validity and reliability, namely the Technology Acceptance Model (Davis et al., 1989) and the Immersive Tendencies Questionnaire (Witmer and Singer, 1998). Students completed the pre-test survey one week prior to the intervention. The pre-test survey consisted of:

- Ten hypoglycaemia multiple choice questions (MCQs) written in conjunction with diabetic nurse specialist nurses (DSNs). The questions were linked to the diabetic case study.
- A series of Likert scale questions (for attitude questions) which were linked to the conceptual framework and hypothesis
- Open-ended questions to stimulate expression of experience and views about using the case studies.

Second Year Adult and Mental Health nursing students at BU undertaking a Long-Term Health Conditions Unit of work were invited to take part in the current study (as they usually receive a diabetes lecture as part of their normative curriculum).

The digital link for the Bristol Online pre-test survey was emailed to the students prior to the experiment day to aid easy access to the survey. Students were sent an explanatory email detailing the study and making them aware that they would be given time to complete the pre-test survey within the lecture one week prior to the study day. Some of the learning from the unit was flipped and made available on the Brightspace VLE to enable time for this to take place. The survey took on average less than eight minutes to complete.

The pre-test survey contained details of the study, in order that participants could make an informed decision about whether they wanted to take part in either the Control or Experimental group. The pre-test survey also contained the ethical consent form. One survey was discounted because the student had made an error in their response to the ethical consent form. It was observed that any student who did not want to take part had received the information in advance and had therefore brought along other work to do whilst the participants completed the online surveys. All students who did not want to be part of the study completed the paper-based exercise for their own learning purposes.

Participants were invited to remain behind after the teaching and learning exercise to complete a post-test, which contained some of the same questions as the pre-test (for comparison), e.g., the hypoglycaemia MCQ questions. This very short space between the intervention and post-test was chosen to limit students encountering any other diabetes learning in the interim. However, it is acknowledged that this means the post-test will measure any immediate surge in knowledge but will not measure any longer-term retention of learning.

4.6.7 Focus Group Discussions

Individual interviews were considered, though FGDs were thought to be more suitable for the qualitative research conducted in this thesis. For example, Kristiansen and Grønkjær (2017) found that the main advantage of focus groups involves how group interactions can reveal both the participants' perceptions, attitudes, and framework of understanding, as well as recognising group norms, sub-cultural and cultural values. The collaborative nature of the focus group offers another advantage. Consistent with applicable literature on the implementation of FGDs (Markova et al., 2007), in the group discussions both debate and expansion on initial responses were fostered, as well as deeper meanings that contextualised the responses gathered via the survey.

Bryman (2009) stated that the FGD technique *"allows the researcher to develop an understanding of why people feel the way they do."* It offers researchers the opportunity to investigate the ways in which individuals collectively make sense of a phenomenon and construct meanings around it. It was also hoped that they would provide evidence relating to the mechanisms and role of agency involved in the teaching scenario (Porter, 2015).

Other advantages of FGDs include that the respondents are given a voice, they provide quick results, and that ideas and themes can be developed more than in individual interviews (Muylaert et al., 2014). In addition, FGDs can produce "collective narratives" on the research issues that progress beyond individual outlooks to create a group viewpoint on the issue discussed, which produces a different level of data from that obtained in individual interviews. For example, Mkandawire-Valhmu et al. (2013) presented a group narrative of “stigma” experienced by women with HIV/AIDS, and an account of women’s “powerlessness” to prevent infection because of societal structures. FGDs can therefore provide a distinctive type of data and viewpoint on the research.

The FGD sought to acquire more information and detect contradictory opinions, if present, in the groups. To make each interviewee feel as comfortable as possible, the interviewer encouraged them to contribute as and when they wanted to do so. However, one limitation of FGDs is that the combined view of the groups is being sought not the individual’s views, though not all academics agree with this criticism. Muylaert et al. (2014) believe that during the group discussion participants share their views, listen to views of others, and possibly refine their own schema after considering what they have heard. As the discussion proceeds contributors begin to ask for clarification from others, which may prompt them to raise additional concerns or share comparable experiences, consequently improving the clarity and richness of the discussion.

Furthermore, volunteering to take part in the FGDs might have appealed to those who viewed the VR intervention more positively, but this would have biased the results. Once again, other academics refute this and claim that participants who have had a negative experience of the intervention might also be keen to air their views within an FGD setting. It is also important to note that FGDs were only held with Experimental Group participants.

A sample of participants from each of the different cohorts was invited to FGDs; LTs also took part in them. FGDs were selected because they draw out opinions shared publicly, which supplemented the individual-level responses attained via the survey. Five FGDs took place immediately after the intervention, with between five and eight participants in each group. The justification for this variety of sample size stems from the ideal that FGDs must comprise enough contributors to yield variety in information provided, yet they should not comprise too many participants because large groups can produce a setting where participants do not feel at ease sharing their experiences.

Using multiple FGDs enables the researcher to evaluate the extent to which saturation has been achieved (Sandelowski, 2001). Morgan (2019) established that three to six different groups are adequate to reach theoretical saturation; the current study included five FGD sessions in total, therefore, it is anticipated that theoretical saturation was achieved. The author acted as a moderator responsible for facilitating the discussion using a semi-structured approach to enable the voice of the participants to emerge. Prompts and cues were used to encourage participants to fully express their views and experiences and to establish an informal discussion context rather than a more formal interview approach. For example, the author asked, “Who has a different perspective on that?” and “How do others feel about that point?” Participants were provided with refreshments, as they had just completed the hour-long diabetes training exercise.

4.6.8 Qualitative Data Analysis Procedures

Following the literature on mixed methods inquiry (Vinson, 2019, Kimmons and Johnstun, 2019, Younas et al., 2019) methods were entirely assimilated during the examination and explanation of results, when results were compared and contrasted from the quantitative and qualitative data sets. The researcher executed simple, exact transcriptions of FGD recordings. Transcripts were imported into the package NVivo Version 11, to code and analyse the qualitative data via thematic analysis (Clarke and Braun, 2018). Both inductive and deductive coding techniques were utilised enabling exploration of the potential mechanisms and influences that had already been identified and to identify new sub-themes emerging from the data.

Theory driven codes that were the starting point were derived from the four CR themes: engagement, confidence, knowledge and immersion. Descriptive quotes were chosen based on the quality and succinctness with which students expressed each point, and on the representativeness of the quote in relation to the general theme. Qualitative questions from the survey were synthesised using thematic analysis (Clarke and Braun, 2018).

4.6.9 Quantitative Data Analysis Procedures

Quantitative data management and descriptive analysis were implemented using Excel, SPSS (22.0) and Smart-PLS 3 software. Excel and SPSS were used to carefully prepare and format the data to be imported into SMART-PLS software; all the statistical analysis was performed in the latter. A 5% significance level was used throughout the evaluation (Corp, 2013). The significance level, also signified as α, is the probability of declining the null hypothesis when it is true. For instance, a significance level of 0.05 shows a 5% risk of concluding that a variance exists when there is no tangible variance (Cohen et al., 2017).

SmartPLS 3.0 software's graphic user interface is intuitive (Sarstedt and Cheah, 2019) enabling users to specify and estimate path models. For instance, by dragging and dropping items from the indicator's menu onto the modelling window, users can easily create new constructs. The constructs are then linked by drawing single-headed arrows among them. This user interface enables the inclusion of quadratic effects (Hair et al., 2018) and moderating effects without having to process data outside SmartPLS 3.0.

SmartPLS is the most complete software for conducting PLS-SEM analyses (Henseler and Hubona, 2016). The SmartPLS 3.0 software enabled evaluation of the quantitative aspects of the survey, harnessing the battery of statistical tests that comprise the PLS-SEM method (Ringle and Sarstedt, 2016). Within SmartPLS, the PLS-SEM algorithm is supplemented by separate bootstrapping routines (Streukens and Leroi-Werelds, 2016), which allow deriving standard errors for all model parameters. The standard results output is user friendly and enables users to browse between descriptive statistics (e.g., indicator and construct correlations) and state-of-the-art measurement as well as structural model evaluation metrics with ease (Henseler and Hubona, 2016, Streukens and Leroi-Werelds, 2016, Sarstedt and Cheah, 2019). As detailed earlier, PLS-SEM allows the approximating of multifaceted cause-effect relationship models containing latent variables (or in CR terms, contextual mechanisms) (Hair et al., 2016).

The SmartPLS 3.0 software was selected for several reasons: ease of use, it is a robust approach incorporating various reliability and validity tests, and finally it is a visual approach that enables researchers (not necessarily from a statistical background) to understand and model abstract statistical concepts. This means that researchers who have dyslexic or dyscalculic tendencies are also able to use this software to analyse the data of complex real-life interventions. Other free software packages also allow some PLS-SEM capabilities that users can revert to free of charge (for example modules in R); however, these are targeted at more experienced users with profound programming skills and come with highly technical documentation. Hair et al.'s (2016) Primer on PLS-SEM is strongly tied to the software and has been translated into various languages. Novice users benefit from such detailed documentation of the software.

The current study is concerned with concepts rather than individual items, hence the suitability of the statistical method chosen, as PLS-SEM enables an expression of the relationships between “collections of ideas” rather than individual constructs. With the support and advice of a BU data analyst it was possible to run the statistics through the entire conceptual pathway model, and more precisely, through the “original pathways” of the models. The aim of this was to provide insight into the big picture (how significant are all the things that happen when you use VR in a lecture theatre?) and evidence to support or reject the hypotheses about the scenario.

4.6.10 Ethical Considerations

Prior to initiating the study, ethical clearance was obtained from Bournemouth University Research Ethics Committee. The ethical review number assigned to this work was 21833. Participants completed a consent form prior to taking part in the intervention and were fully informed about data protection and anonymity. The project was low risk. Participants were informed that their participation was voluntary and that they could remove their consent and data from the study at any stage. Confidentiality and data protection measures were implemented during this study as recommended by BU. Potential participants were provided with the contact details of members of the research team, whom they could contact if they required any further details or clarification. All the qualitative interviews with students, LTs and an academic were recorded and transcribed, and any identifiable information was removed from the transcriptions. All participants received a pseudonym to guarantee that when accumulating, storing, and reporting qualitative findings, it was not possible to identify individual participants. All data were stored on password-protected computers and according to institutional regulations.

An amendment was sought (and granted) as ethical clearance was requested to extend to 2nd Year Mental Health nursing students in addition to Adult nursing students. The reason behind this was because they receive the same diabetes lecture in their second year of teaching. It was thought that it would be more ethical to enable both cohorts (Mental Health and Adult) to have the opportunity to take part in the study. The online ethics checklist has been included as Appendix Eight.

Technology and software that was utilised for the teaching and learning exercise was tested (by a group of eight LTs) prior to the study taking place, to ensure that it would work regardless of variation in desktop devices being used by the students. LTs were present on the day to help troubleshoot any technological issues. Advice was sought (from the unit leader) to ensure optimal recruitment of research participants (e.g., running the consent session on their induction day, and so forth). Careful planning involved liaison with faculty staff to optimise the chances that there would be enough time for participants to successfully complete the tasks using the technology and that FGDs could be of adequate length.

4.7 Chapter Summary

In this chapter the research paradigm was described along with both the methodology and methods underpinning the empirical study. Justification for the method and methodology were also explained and justified, including reasoning behind selecting an RCT approach, analysed via PLS-SEM, and underpinned by a CR angle. Aims, objectives and hypotheses were all detailed. Rationale was provided for the discipline focus. A description of how the VR exercise developed was delineated. Ethical and risk assessment considerations were described. Data collection and qualitative and quantitative analysis approaches were outlined. The next two chapters will report first the findings of the quantitative results (Chapter Five); followed by reporting of the qualitative results (Chapter Six). The CR angle will be interwoven between both these results chapters, the Discussion and Conclusion.

Chapter 5 Reporting of Quantitative Results

5.1 Synopsis

This chapter reports quantitative findings including pre and post-test survey data. Research aims, questions and hypothesis are reiterated. The demographics of the Experimental and Control Groups are also reported. The initial part of the chapter focuses on the Partial Least Squares-Structural Equation Modelling (PLS-SEM) pathway model, describing how it was tested for validity and reliability prior to running the pre and post-test survey data through the pathway model. The six stages involved in this validity and reliability checking of the PLS-SEM process, are explained in depth. Following explanation of the methods used to test the pathway model, the PLS-SEM findings are reported.

5.2 Quantitative Methods

The current study aimed to compare the difference between VR learning and traditional methods via an RCT from a CR angle. The RCT consisted of 2nd year Adult and Mental Health nursing students at the Bournemouth and Yeovil campuses completing a pre-test survey, a Control or Experimental diabetes case study, and a post-test.

Statistical significance was sought to evidence whether the desktop VR provided enhanced student learning gain in this situation. The pairing of CR with PLS-SEM also enabled evaluation of the individual variance between students' experiences when learning in a style or way, in this case using VR. Linking back to Chapter Three, the pairing of CR with PLS-SEM also enabled the testing of the proposed CPMAO configurations that might interplay when HE students learn using desktop VR. Two CPMAO configurations were proposed, drawing upon the CR literature as follows:

Table 5-1 Recap of the CPMAO Configurations tested in this thesis

CPMAO configurations	Contextual Mechanisms	Programme Mechanisms	Agents	Outcome
CPMAO 1#	Amount of diabetic nursing experience and relevant computing EXPERIENCE Required to complete exercise and understand complex diabetic concepts	Affordances of VR including opportunities to practise via experiential learning, and affordances of computing exercise, e.g., immediate feedback given and the chance to repeat the exercise	CONFIDENCE boosting via feedback	Improved KNOWLEDGE of complex diabetic issues
CPMAO 2#	Participant ENGAGEMENT level with the learning content	Affordances of VR, e.g., visualisation of abstract concepts	Participant IMMERSION with the desktop VR	ENJOYMENT of the learning activity

Five hypotheses were developed from the CR literature review. These hypotheses were designed to be tested via the CR/PLS-SEM pathway analysis and are as follows:

H1a: experience will have a positive influence on knowledge

H1b: confidence will mediate the experience to knowledge relationship

H2a**:** engagement will have a positive influence on knowledge

H2b: immersion will mediate the engagement to knowledge relationship

H3: the experimental group will be more knowledgeable than the control group

5.3 Survey Demographics

In total 171 students completed both the pre and post-test surveys. This was a response rate of 67%. According to Fincham (2008) a response rate of approximating 60% should be the goal of researchers and is expected by Journal Editors; following this academic convention, the response rate for the current study was deemed acceptable. However, it was noted that the number of students attending the lecturers was low, though of those who did attend most completed the survey (81%).

There were no missing data from the surveys because the "complete all" option was set on the BOS software disseminated to the students. Paper copies of the survey were made available to students in case they did not have a computing device with them; however, all students completed the survey digitally. The demographics of the survey respondents are displayed in Figure 5-1

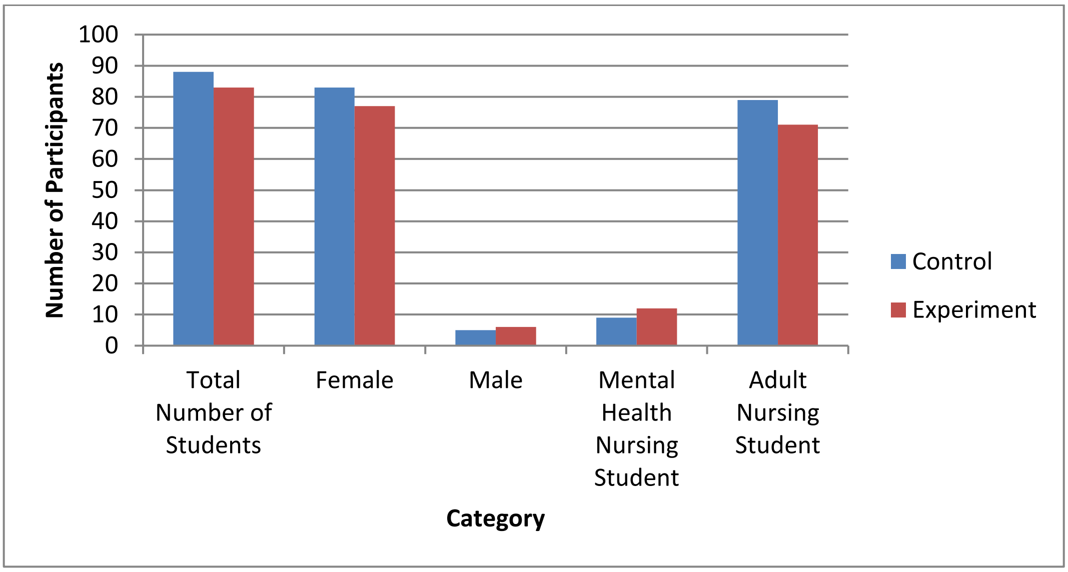


Figure 5-1 Survey Demographics

The data show that the number of respondents in the Control (n=88) and Experimental group (n=83) were comparable. The number in each group was sufficient according to Cohen’s (1992 cited in Hair et al., 2018) PLS-SEM sample size recommendations, which had stated that to detect an R^2 value of at least 0.25, 45 participants would be needed in each group to obtain an 80% statistical power, with a 5% probability of error. More females took part in the study than males (only 6% in total). This is representative of both cohorts at BU and of the nursing workforce nationally, as only 11% of the NHS workforce is male. Finally, Figure 5-1 illustrates that more Adult nursing students took part in the study than Mental Health nursing students; once again this is due to a smaller cohort of the latter. Overall, the Control and Experimental Groups were comparable in terms of gender and type of nursing being studied.

Participants were also asked what their age was. When age was tested with an Independent Samples Median test across Control and Experimental Groups, there was found to be no difference between groups (*P= .118*), thus the ages of students in both groups was comparable. As can be seen in Figure 5-2, one third of the respondents were in the 21-25 age brackets. Many students were younger than 20 years of age (29%). Very few students were over the age of 40 (approximately 8%).

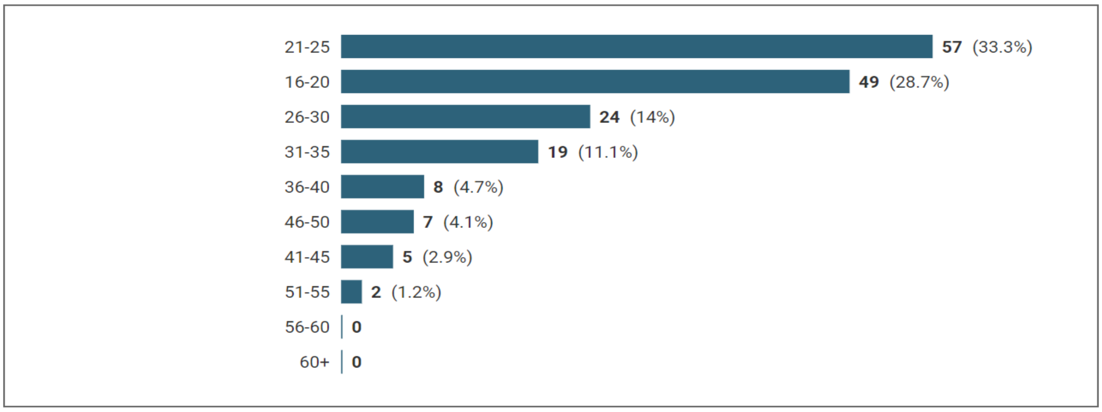


Figure 5-2 Respondent Demographics- Age- taken from the Bristol Online Survey

Additionally, respondents were asked to detail any relevant qualifications that they had obtained (Q22 shown in Figure 5-3).

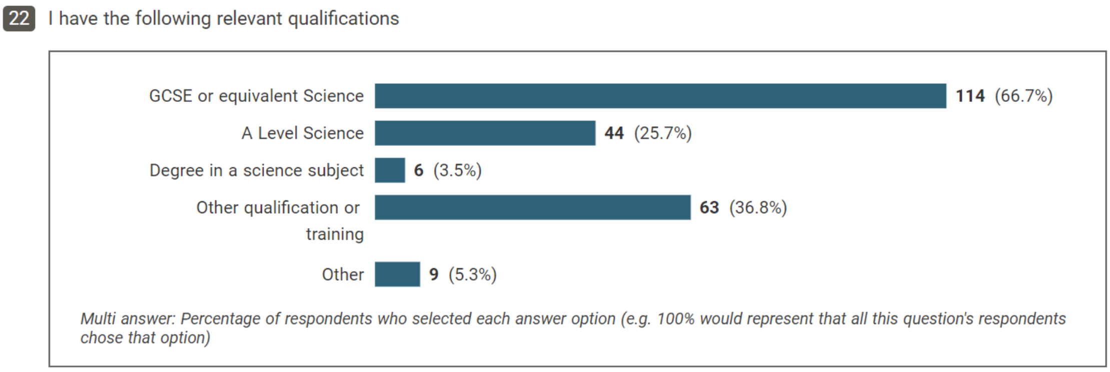


Figure 5-3 Participant Demographics- Relevant Qualifications, showing that less than a third of students had obtained an A Level in a Science Subject.

Over one third (67%) of students have a GCSE (or equivalent) in science, whilst under one third (26%) have obtained an A Level in a science subject. Very few (4%) have a Science Degree. The categories for other (42% in total) included GNVQs, e.g., in Health and Social Care. This might have had an impact on students' knowledge and understanding of hypoglycaemia, which is, anecdotally, deemed to be a complex and abstract physiological concept.

Prior diabetic nursing experience was sought in Question 20 of the pre-test survey. Figure 5-4 reports the results of that question.

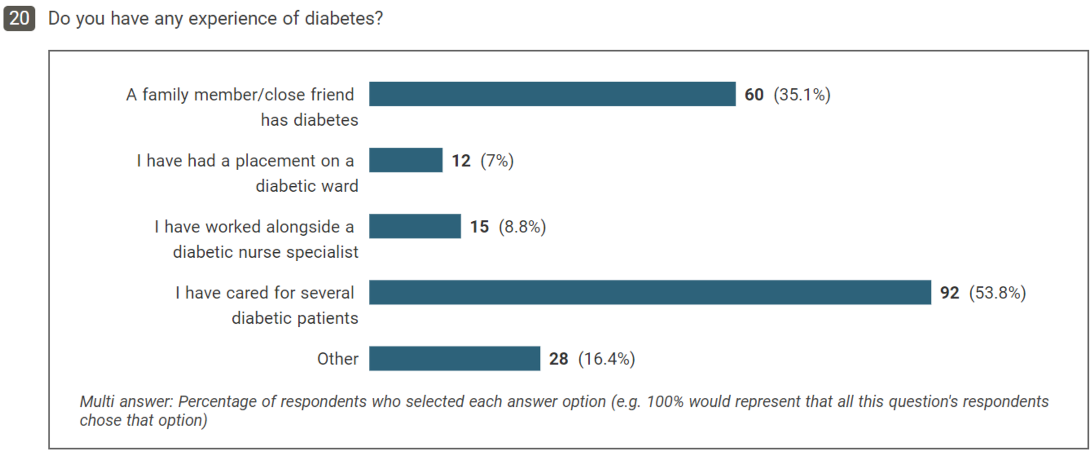


Figure 5-4 Students' prior diabetic nursing experience, showing that over half of the student had cared for diabetic patients.

Over half (54%) of students had previously cared for a patient/patients with diabetes, whilst just over one third (35%) had a family member or close friend who has the disease (in some instances the student themselves had diabetes and this was reported in the other section). Few students had either worked with a Diabetic Nurse Specialist (9%) or had had a placement on a dedicated diabetic ward (7%).

Both Yeovil and Bournemouth cohorts were invited to take part in the study. Figure 5-5 shows the distribution of each cohort across the Groups.

12 Which cohort are you in?

	Your survey (83 responses)		Comparison (88 responses)		Total (171 responses)	
Option	Raw	%	Raw	%	Raw	%
Bournemouth Based	75	90.36%	77	87.50%	152	88.89%
Yeovil Based	8	9.64%	11	12.50%	19	11.11%

Figure 5-5 Respondent Demographics- Cohort- taken from the Bristol Online Survey

The data show that the randomisation between groups was equal between the cohorts. In both cases more students were randomised into the Control Group than the Experimental Group. Having a slightly higher sample sized Control group is advocated by (Riniolo, 1999).

5.4 Pre-test Survey Findings

216 (81% of the total population) students completed the pre-test survey. Part of the pre-test survey included ten MCQs relating to diabetes and hypoglycaemia. Whilst 85% (n=145) of students could confirm a diagnosis of hypoglycaemia, only 15% (n=28) correctly cited the signs and symptoms of the condition (Figure 5-6).

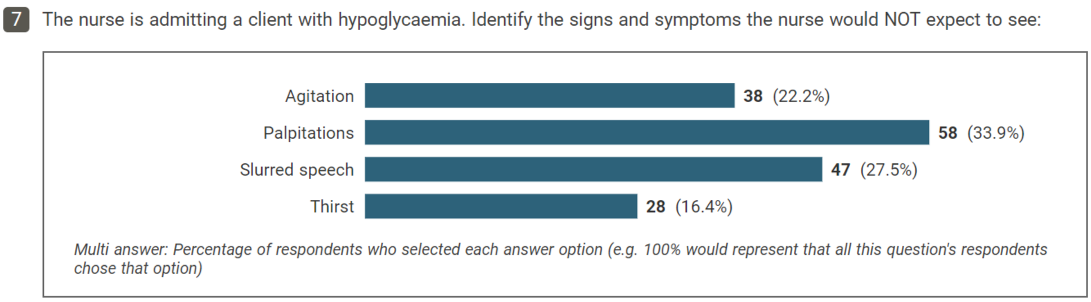


Figure 5-6 Results of the question relating to the "Signs and Symptoms not associated with Hypoglycaemia". The figure shows that the students were unsure about the signs and symptoms of Hypoglycaemia.

Only 29% (n=50) of students confirmed that a sulphonyl urea could cause hypoglycaemia.

It can be seen from the data in Figure 5-7 that only 8% (n=13) would provide an additional dose of glucose if the patient's blood glucose were still below 4mmol/L after 10-15 minutes.

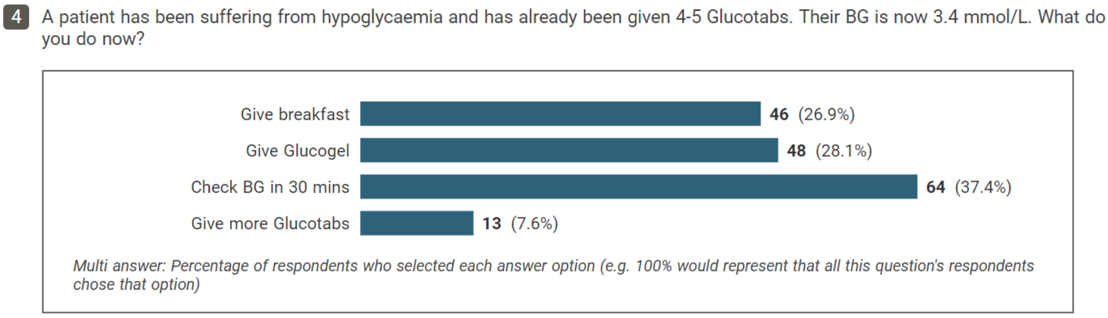


Figure 5-7 Clinical Decision-making around Glucotab administration demonstrating a lack of knowledge about the correct treatment for hypoglycaemia.

Only 17% (n=28) of students considered that their knowledge of hypoglycaemia is good, however, 98% (n=167) stated they were interested in learning more about diabetes. Students were asked to rate how often they used various digital devices (Question 14). The results are shown in Figure 5-8.

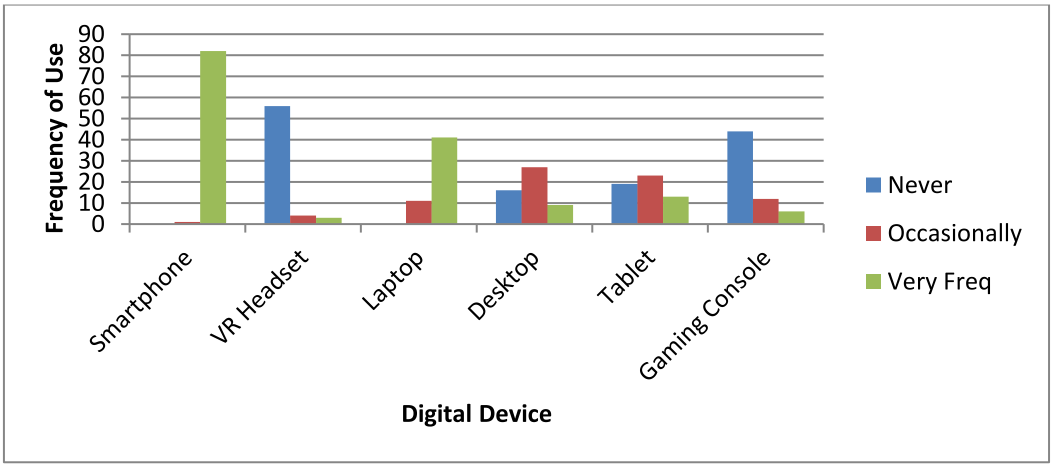


Figure 5-8 Student Digital Device Usage showing frequent use of smartphones

A notable finding from these data is that Smartphone usage amongst the students was high, with over 80% reporting very frequent usage. Laptop devices were used frequently by less than half of the students. Second Year Nursing students' usage of VR headsets and gaming consoles was generally very low; this is representative of similar samples where the cohorts are largely female (Washington et al., 2019). Students were also asked how often they performed certain digital activities using their computing devices; see Figure 5-9 (Question 15).

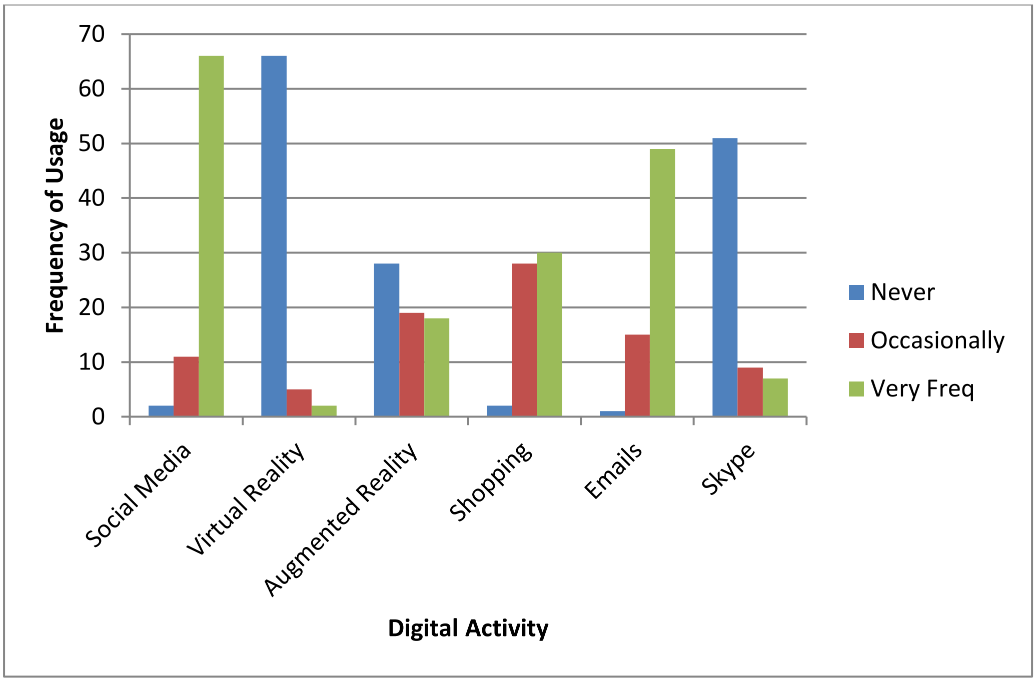


Figure 5-9 Student Digital Activity Frequency demonstrating frequent use of social media

Data shown in Figure 5.9 reveal that most (approximately 65%) Second Year Nursing Students use their computing devices to interact with social media on a very frequent basis. Emailing was also an activity that almost half of students undertake using their digital devices. Virtual Reality and Skyping were not very popular activities amongst students, as is evident from the data. An unexpected finding was that over one quarter of students reported very frequent use of AR (e.g., AR elements of apps including Snapchat). This might be because AR is becoming available in a variety of easy to use social media apps.

In Question 21 of the pre-test survey, students were asked to rate their prior knowledge. The results of this question have been presented in Figure 5-10.

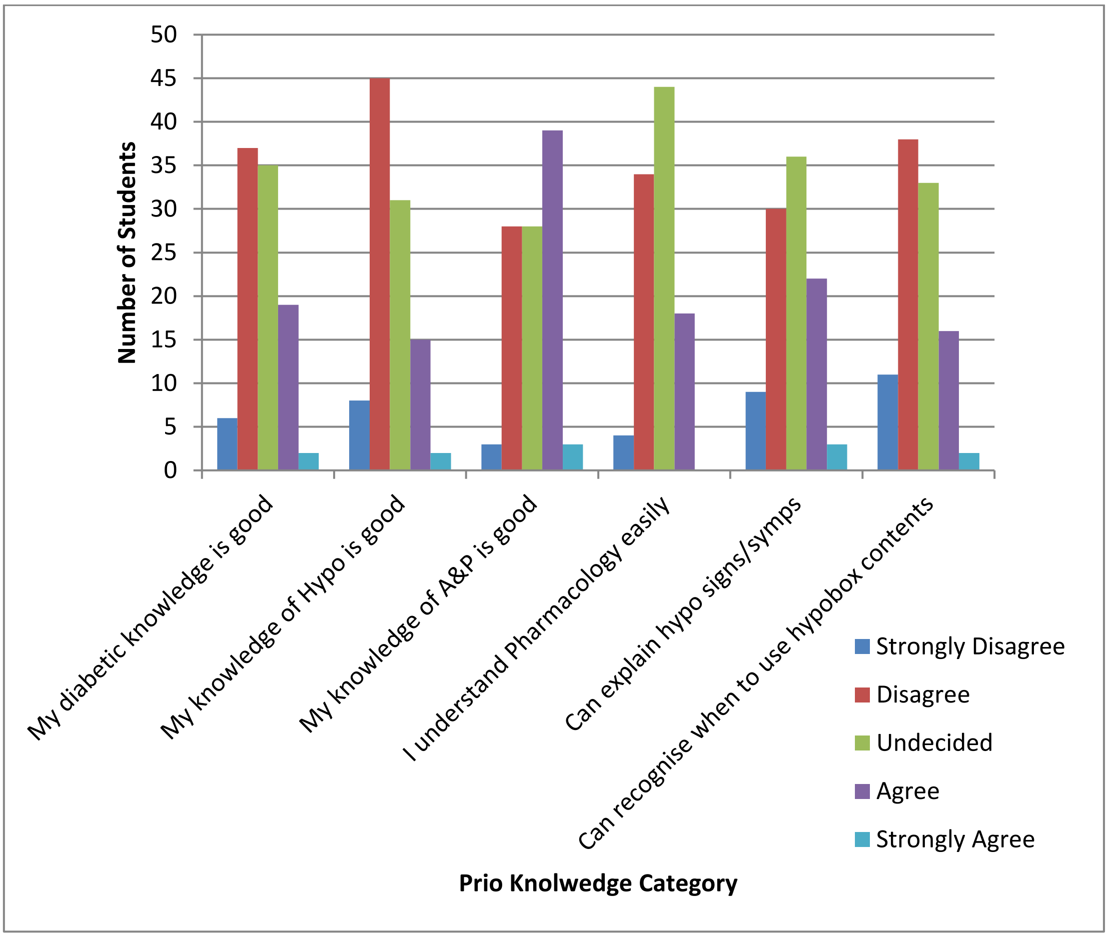


Figure 5-10 Students' confidence in their nursing knowledge prior to the intervention showing that they felt their knowledge about hypoglycaemia was lacking

The data evidence that the three areas in which students doubted their knowledge were diabetes, hypoglycaemia and when to use each item in the Hypobox. Students were most confident about their knowledge of anatomy and physiology. Students were undecided (approximately 45%) about their knowledge of pharmacology.

5.5 Post-Test Survey Results

Figure 5-11 compares the Control and Experimental group post-test survey results with those of the pre-test survey. As can be seen, both groups improved their hypoglycaemia quiz results after they had completed the case study exercise.

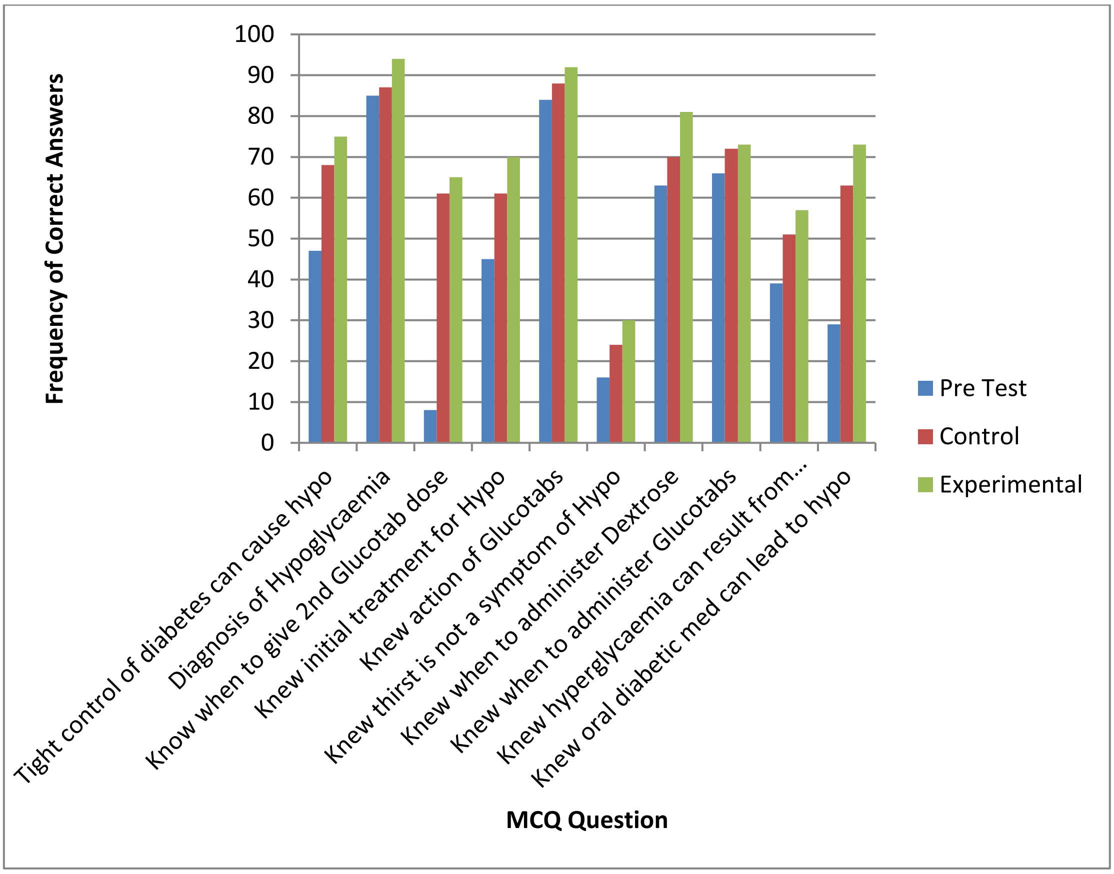


Figure 5-11 Results of the ten Hypoglycaemia MCQs from the Post-test survey demonstrating the efficacy of the VR Simulation

The Experimental group performed better on every question. Figure 5-11 shows that students had much better knowledge of when to administer a second dose of Glucotabs after they had completed the training exercise, with a 61% (n=54) improvement for students in the Control group and a 65% (n=54) improvement for Experimental group students. What is striking about the frequencies in this figure is that the learning surrounding thirst not being a symptom of hypoglycaemia did not improve so dramatically; the training exercises did not focus on this part of the learning but did let students know what the actual symptoms were. The question in the survey did not reflect student learning about thirst; however, it is possible that the learning still took place.

There was an 11% difference between Control and Experimental Groups in terms of learning about when to administer Dextrose, and a 10% difference between the groups surrounding which oral medication can lead to hypoglycaemia. The Experimental group scores were statistically significantly higher ($P<.001$) than those of the Control group at the 0.05 significance level, using an independent sample Median Test. Thus, the Null Hypothesis that medians of Quiz scores are the same across categories of groups can be rejected. That is to say, the Experimental group answered the post-test hypoglycaemia MCQs more efficiently, which is suggestive of short-term learning gain superiority in the desktop VR group.

Students were asked to rate their confidence surrounding their nursing knowledge having completed the VR or paper-based case study intervention. The results are compared in Figure 5-12.

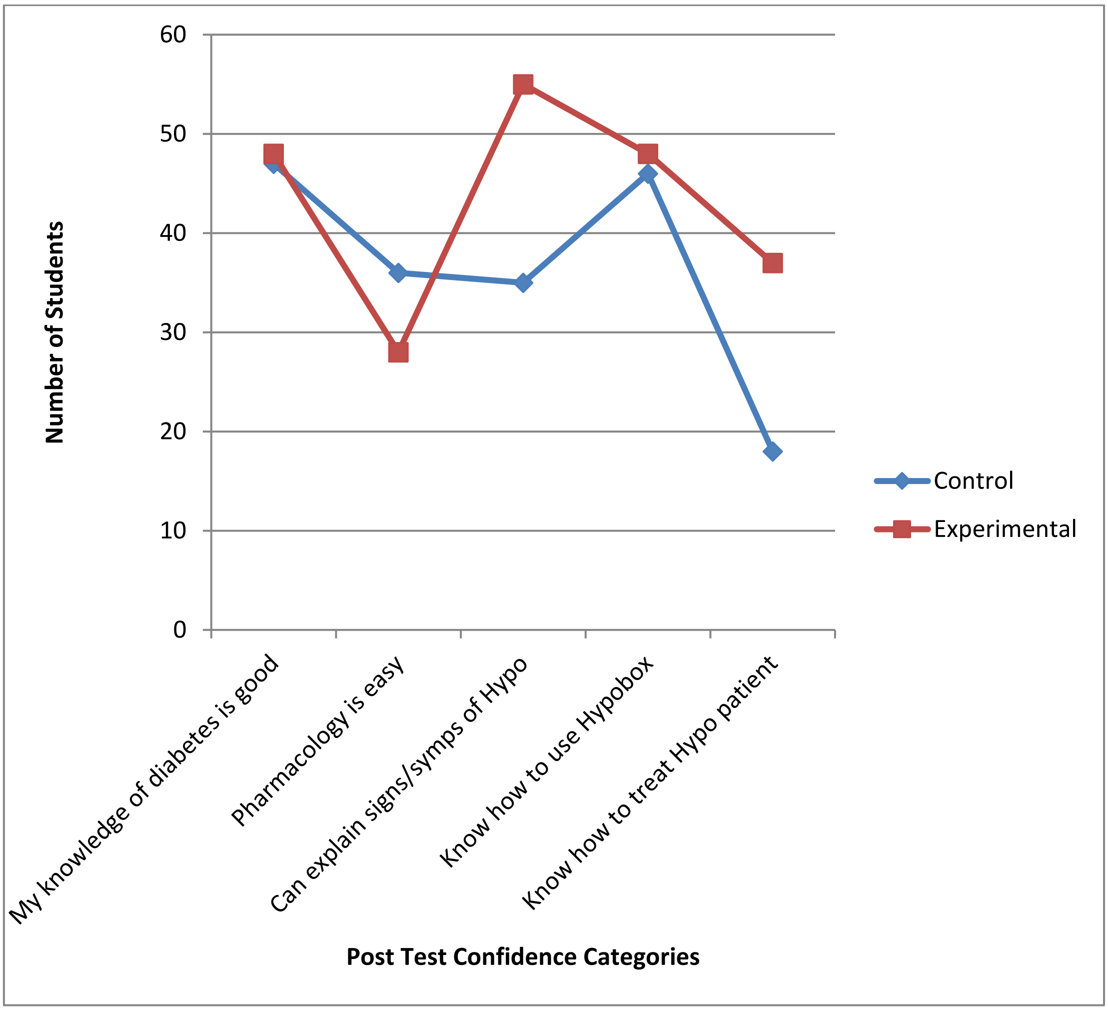


Figure 5-12 Post-test results of students' confidence concerning nursing knowledge

The main observation from the data is that Experimental Group students felt more confident about explaining the signs and symptoms of hypoglycaemia and about how to treat a patient suffering from acute hypoglycaemia than Control Group students. Student confidence about overall diabetes knowledge was comparable. Experimental Group students felt slightly more confident than Control Group students about how to use each item in the Hypobox. An unexpected result was that students felt less confident about their pharmacology knowledge having used the VR-training simulation. This might have been because the VR intervention had highlighted the limitations in their knowledge.

Figure 5-13 displays comparative results of the Pre-test and Post-Test (Control/Experimental Group) student confidence levels concerning their nursing knowledge.

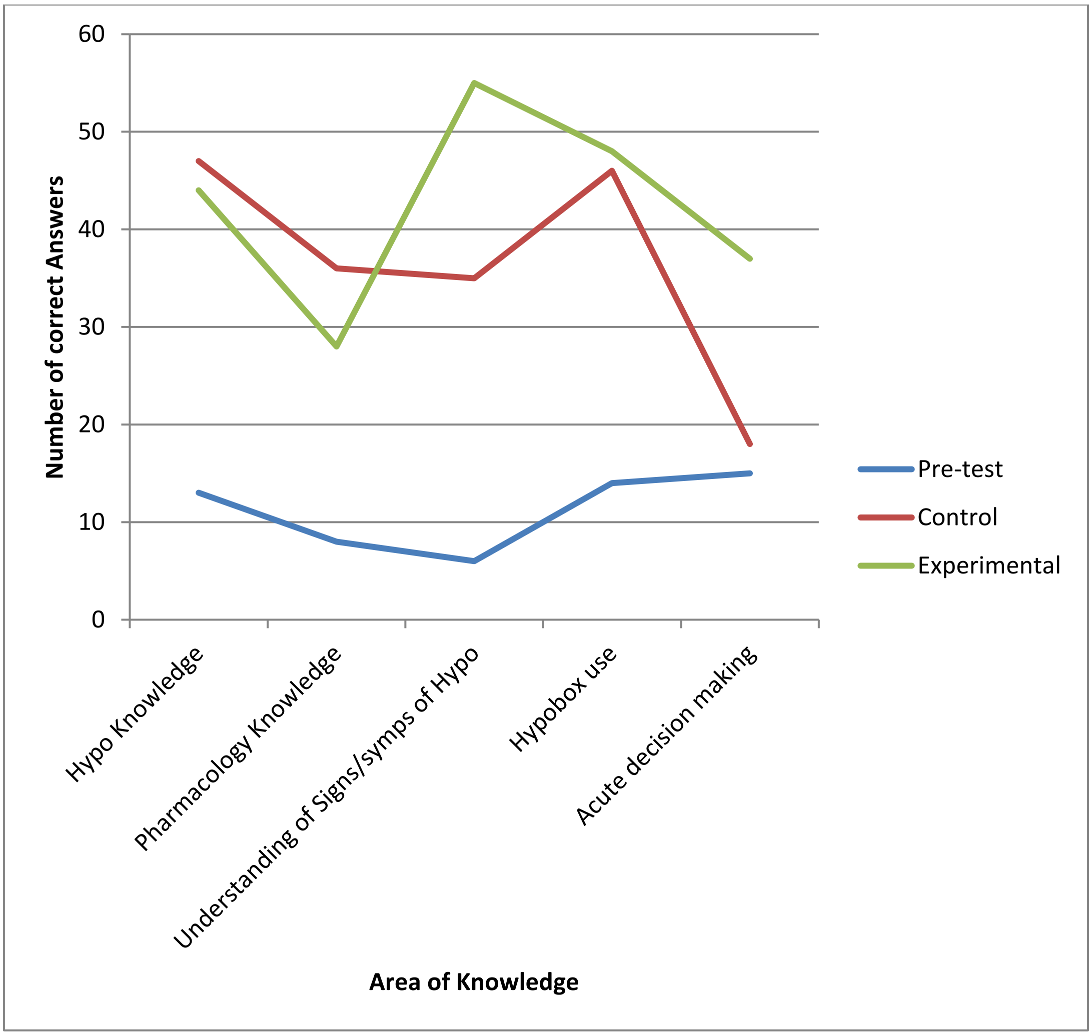


Figure 5-13 Comparison of Short-Term Knowledge Gain between Control and Experimental Groups

The data indicate that both groups' confidence improved post intervention. The Control Group felt more confident about both their general hypoglycaemia knowledge and pharmacology knowledge having used the paper-based case study. The Experimental Group felt more confident about more specific areas of nursing knowledge including how to use each item in the Hypobox. There were two main areas where the Experimental Group post-test confidence gain was much higher than the Control Group and these were in their understanding of the signs and symptoms of hypoglycaemia, and their acute decision-making (knowing how to treat a patient experiencing acute hypoglycaemia) confidence.

Figure 5.14 presents a comparison of short-term nursing knowledge gain of each post-test Group (Control/Experimental) with the pre-test scores. The scores comprise the results of the ten MCQs about nursing knowledge.

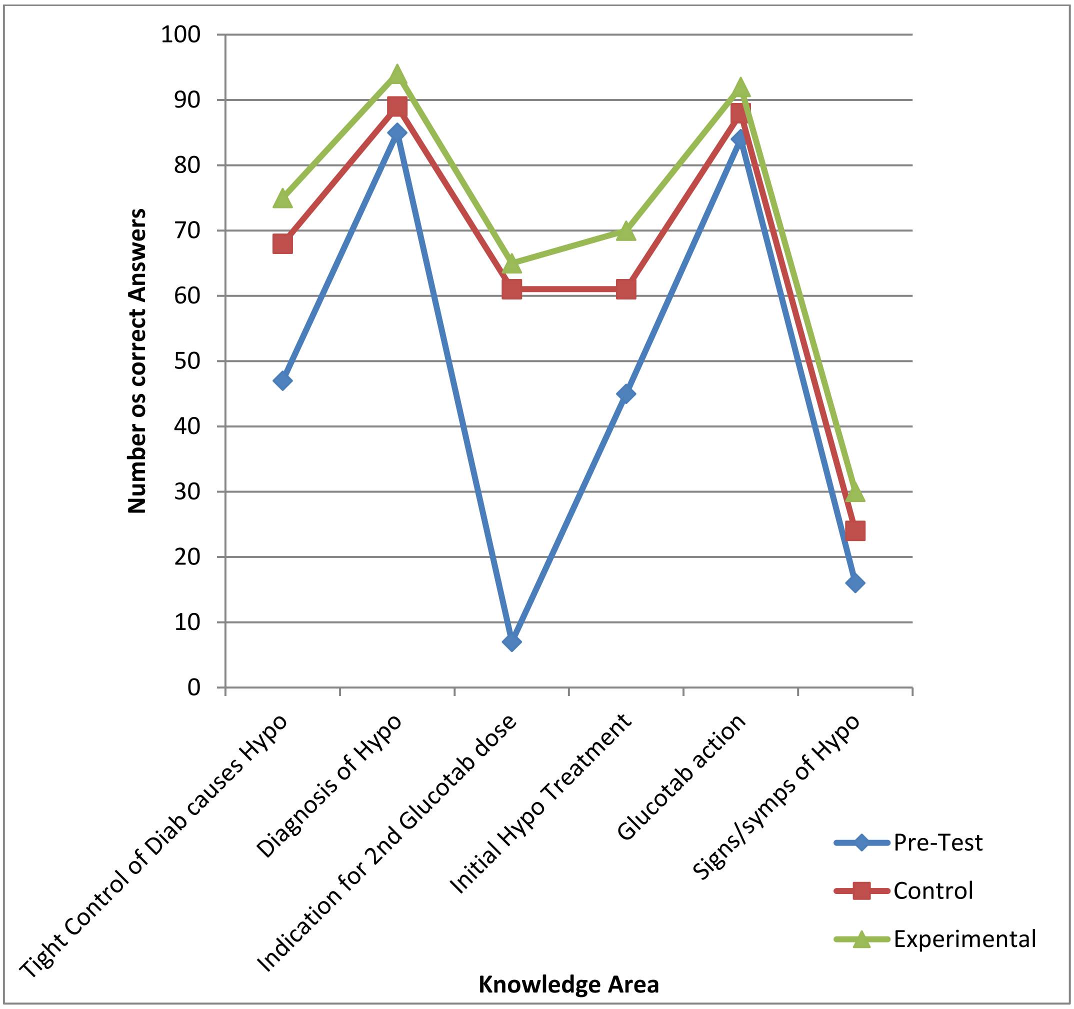


Figure 5-14 Comparison of Nursing Knowledge Pre-Test/Post-Test

The data are interesting in several ways. Firstly, both groups demonstrated more than a 50% learning gain surrounding when to administer a further dose of Glucotabs to a hypoglycaemic patient. Both groups also improved their short-term learning gain (by approximately 20%) concerning the understanding that tight control of diabetes can lead to hypoglycaemia and about the initial treatment recommended for hypoglycaemia (over a 15%

increase in each group). Short-term learning gain did not improve as drastically in the following areas: diagnosis and ability to identify signs and symptoms of hypoglycaemia, and the action of Glucotabs. Overall, the Experimental Group demonstrated superior learning gain when compared to the Control Group. The significance of this gain will be statistically analysed within the PLS-SEM section further on in the thesis.

Figure 5-15 shows a comparison between Experimental and Control groups' learning gain (as measured with the ten MCQs) with their self-assessed Confidence Gain. Results of questions 3, 5 and 7 were combined for the Hypo Knowledge gain scores. Results of questions 6, 8 and 9 were combined for the Hypobox Knowledge gain scores.

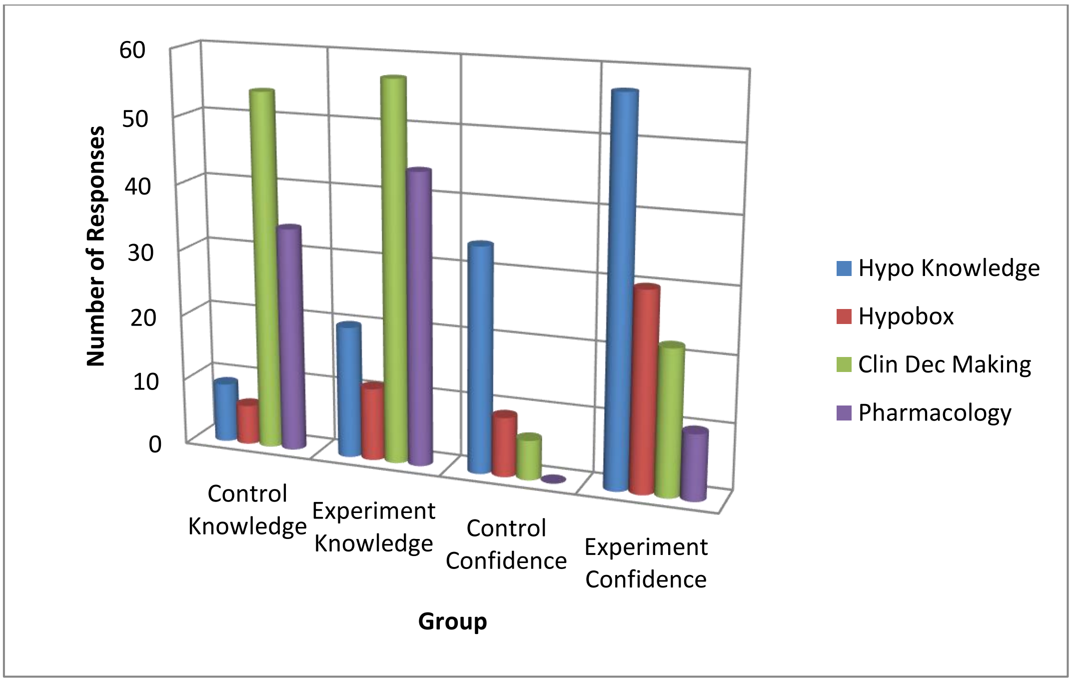


Figure 5-15 Comparison of Short-Term Learning Gain with Confidence Gain

These data evidence that overall short-term knowledge gain and confidence gain were higher in the Experimental Group. Confidence gains were not directly related to knowledge gain. For example, the data show that Experimental Group self-assessed confidence increased by almost 30%; however, the actual measured knowledge gain was only 10%.

Having completed the training exercise, respondents were asked the following question: “Which approach to learning from a diabetes case study do you think is more enjoyable/motivating?” Figure 5-16 presents the comparison between groups concerning their answers to this question.

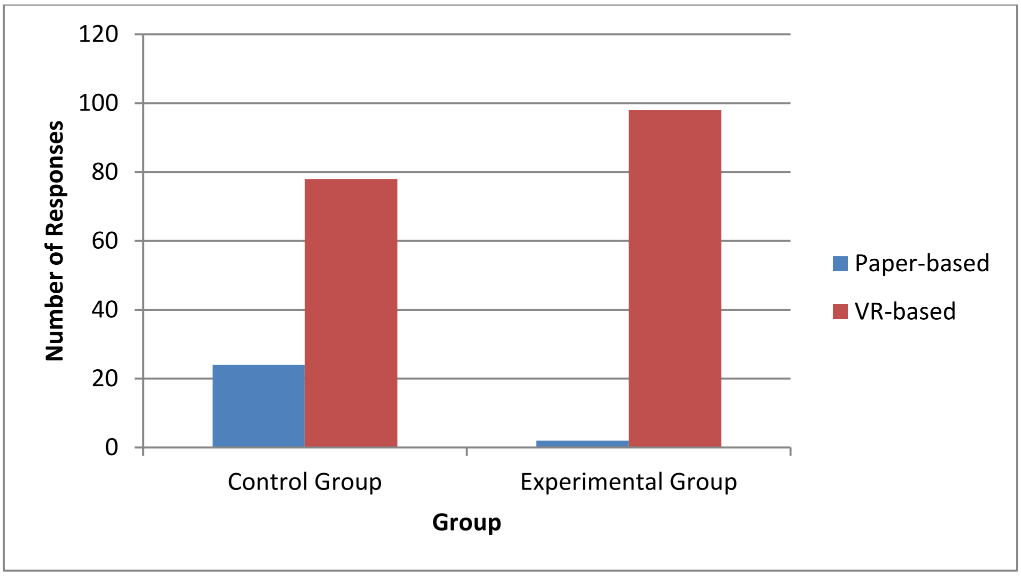


Figure 5-16 Response to the Question- “Which approach to learning from a diabetes case study do you think is more enjoyable/motivating?”

The data show that both groups believe that VR-based learning would be more enjoyable. The participants, who had just used desk-top VR, cited VR as their answer in 20% more cases.

Linked to this were the results from Question 33, which prompted students to rate their interest in various areas. Figure 5-17 shows the findings from both groups.

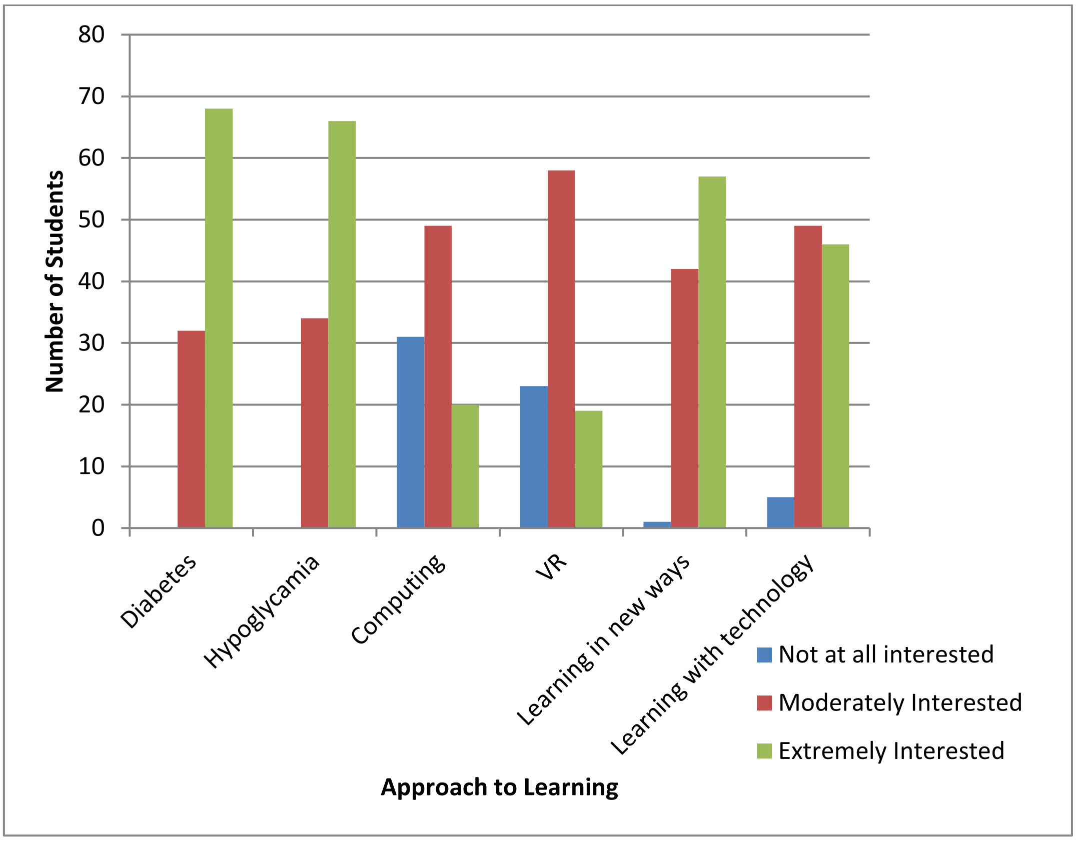


Figure 5-17 Levels of Student Interest in Learning Showing a Willingness to Learn using New Approaches

The data evidence that computing and VR are the ways of learning that the participants (n=171) are least interested in. Diabetes, hypoglycaemia and new ways to learn interested the students the most. However, when a comparison is drawn between groups, the results are more revealing (Figure 5-18).

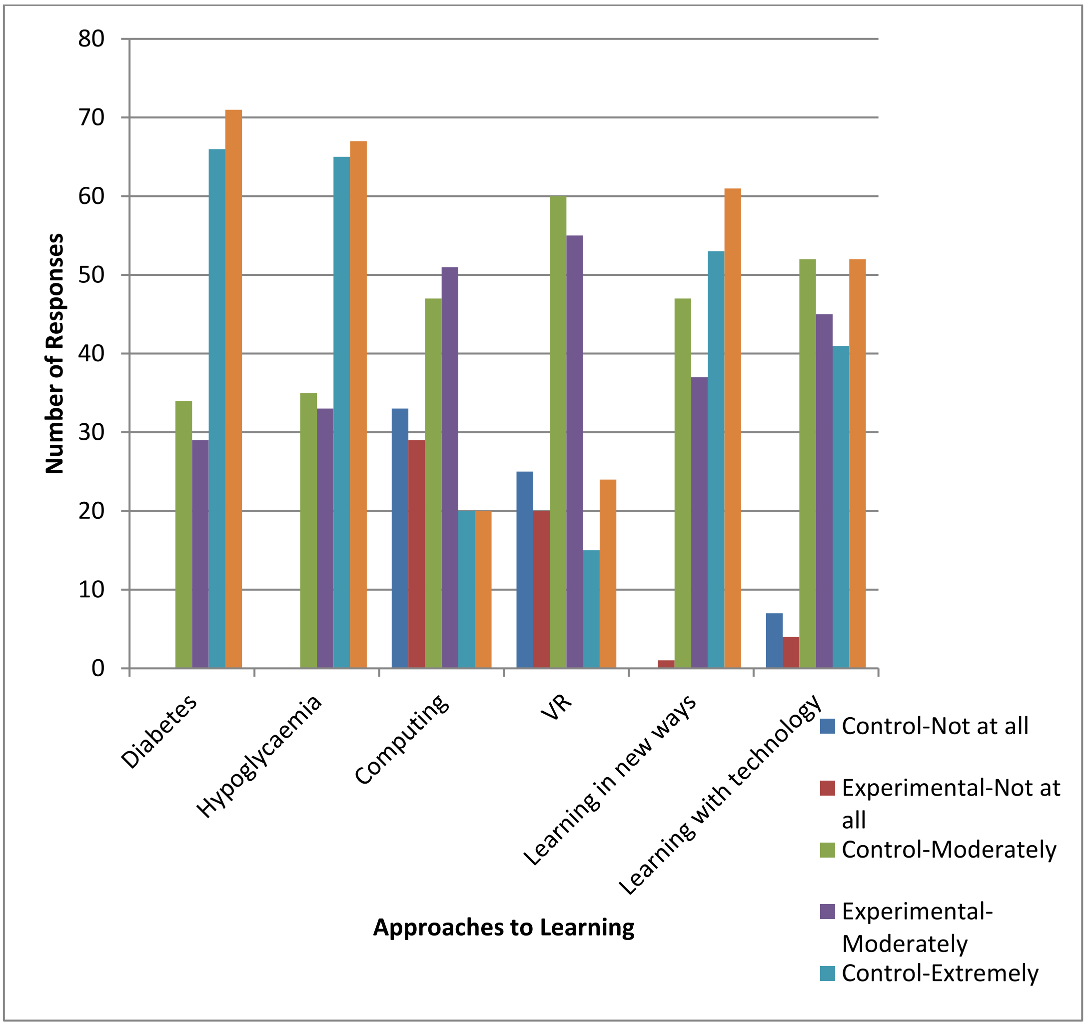


Figure 5-18 Group comparison of student interest in learning, showing that once students had used VR they were more Interested in using it in the future

As can be seen in the data, the Experimental Group were more interested in learning VR than the Control Group. The intervention also sparked their interest (albeit temporarily) in learning in new ways and with new technologies. This finding has ramifications in terms of policy and practice, because it suggests that once students have tried a VR intervention, they are more likely to be open and responsive to other similar TEL opportunities.

Respondents were also asked the following question: “Q27. Which approach to learning from a diabetes case study do you think is the most effective?” The answers can be seen in Figure 5-19.

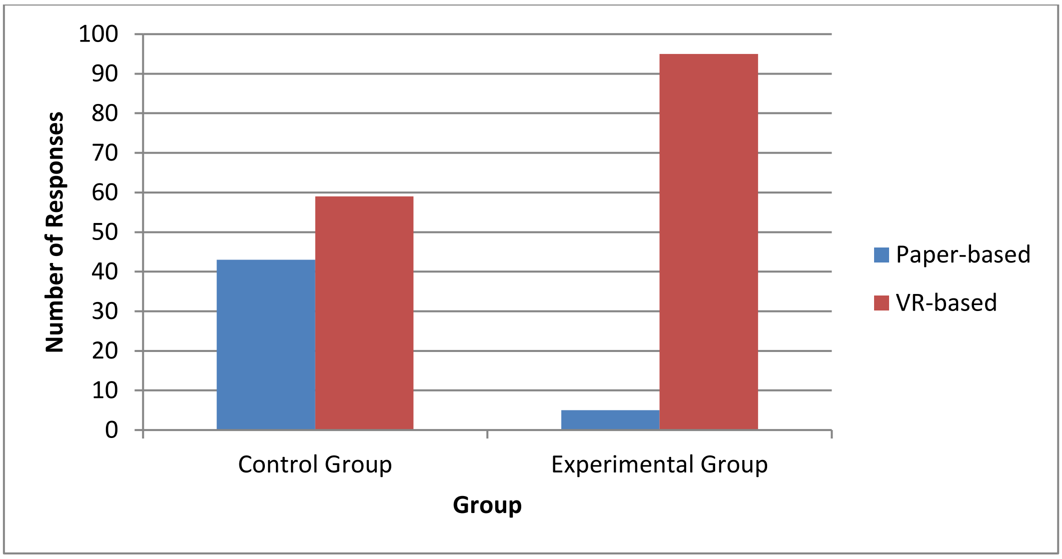


Figure 5-19 Student responses to the question: “Which approach to learning from a diabetes case study do you think is the most effective?” Indicating that once students had used VR they could see the advantages of it for their learning.

The Experimental Group overwhelmingly (95%) thought that VR would be the superior way to learn about complex concepts including diabetes. However, the Control Group (who had not tried the VR diabetes, though had all used VR Dementia Walk Through in a previous session) were more balanced in their responses, with only 16% more learners believing that VR-based learning would be more efficient than learning via a paper-based case study.

In Question 17, students were asked how easy they perceived the VR simulation to be. Results are compared in Figure 5-20 (which shows responses from the Experimental Group).

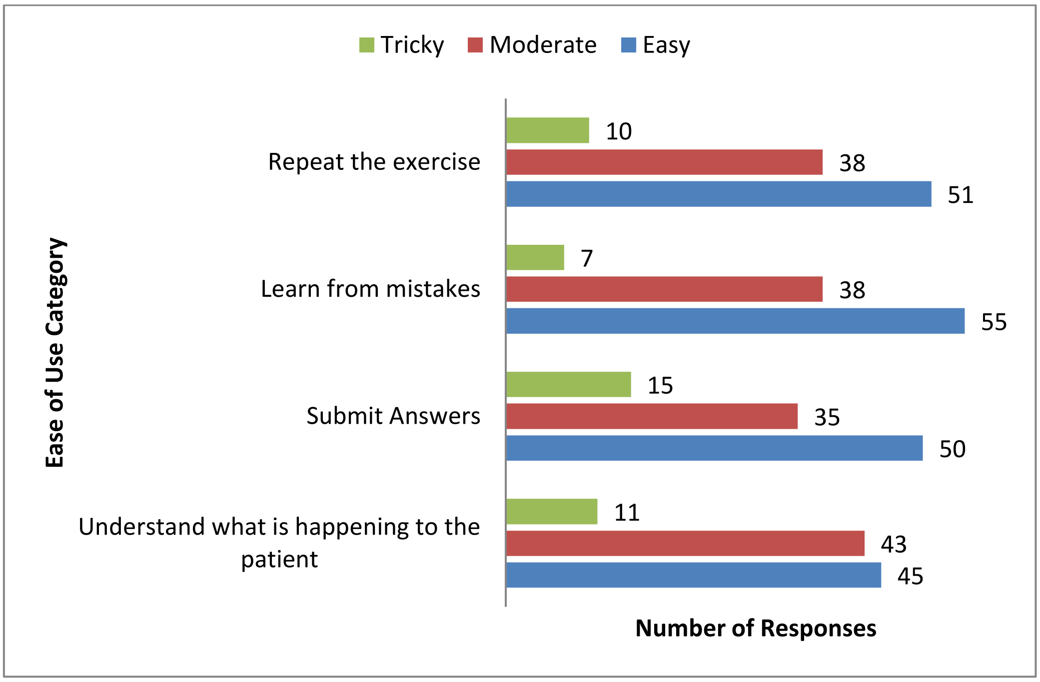


Figure 5-20 Students' Perceived Ease of Use of the VR Simulation, indicating that students found the VR simulation easy to use.

Overall, the answers were wide-ranging, but most respondents viewed the VR simulation as either easy or moderately easy to use. Ease of use scores were highest for the ability to "learn from mistakes" and "repeat the exercise", with 55% and 51% of respondents scoring that way, respectively (multiple selection was possible). "Submitting answers" and "understanding what is happening to the patient" were viewed as being trickier by 15% and 11% respectively.

As part of Question 23, students reported how useful they thought the VR simulation would be for different situations and the results have been displayed in Figure 5-21 (which shows responses from the Experimental Group). Students perceived that the VR learning approach would be useful for the areas of recall, comprehension and application of learning of diabetic concepts and the two latter areas.

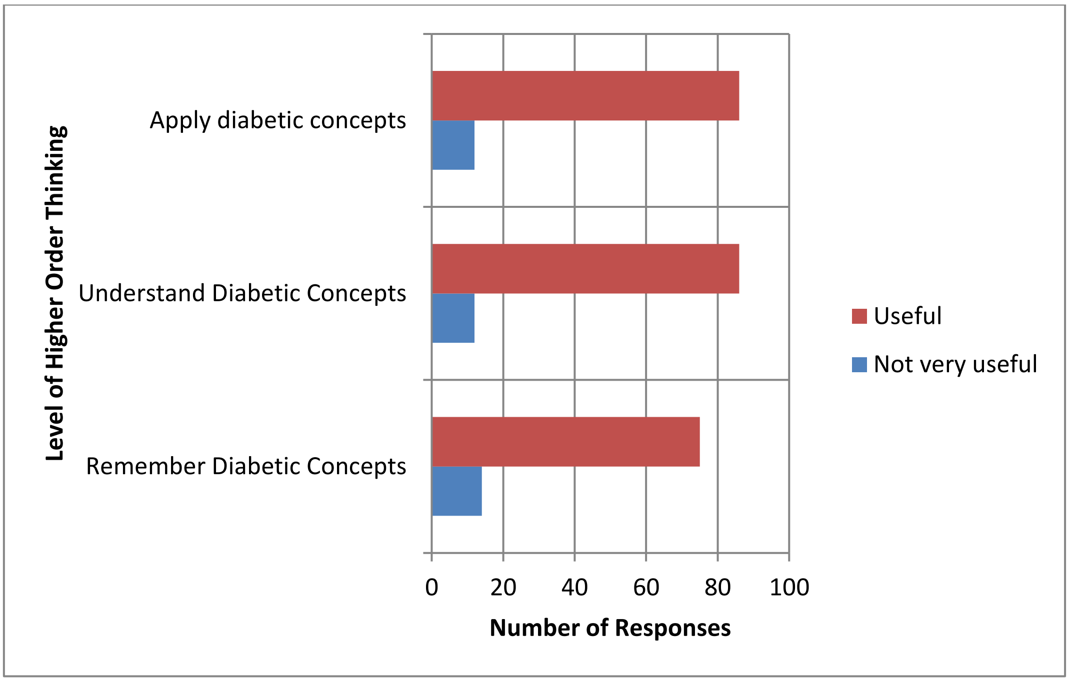


Figure 5-21 Students' perceptions of the Usefulness of the Desktop VR Simulation, indicating that VR simulation could help to apply nursing concepts to clinical practice

Question 25 asked students to rate how useful the simulation would be for use at home and at university. In total 55% of students felt that they would probably use the VR simulation at home and 36% felt they would use it.

Finally, students were asked to rate VR, paper and skills lab (used in other sessions though not in this RCT) and case study learning on a scale of least to most effective (out of 3).

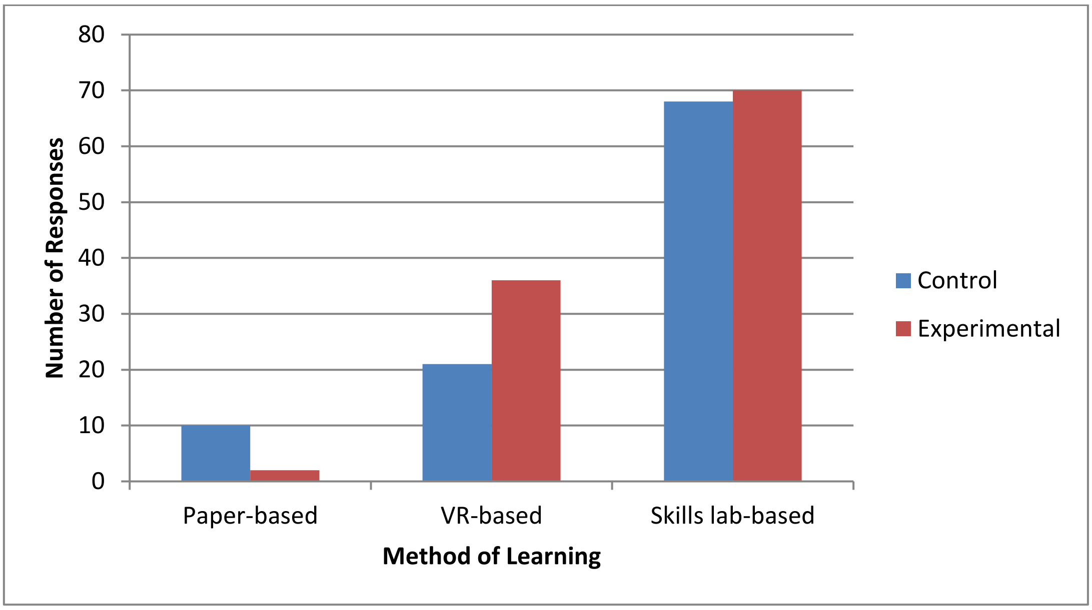


Figure 5-22 Student's responses to the question-"Which approach do you think would help you learn about diabetes most effectively?" Results confirmed that students see VR learning as a bridge between theory and practice.

The main observation from Figure 5-22 is that approximately 70% of students (regardless of which group they were in) felt that learning about a diabetes case study in a skills lab would be the most effective approach to learning. VR learning was deemed to be the second most effective way, followed by paper-based learning.

5.6 The Partial Least Squares-Structural Equation Modelling (PLS-SEM) Process

Prior to running the survey data through the PLS-SEM pathway model using SMART-PLS 3.0 software, there were several steps to complete in order to ensure that the model was valid and reliable. Before proceeding, the definitions of causality (in the context of PLS-SEM) and of reflective and formative measurement are provided. This thesis adhered to the set of six stages that Hair et al. (2016) set out, in order to test that a given model or pathway is ready (valid and reliable enough) to proceed to statistical analysis.

PLS-SEM definitions for formative and reflective measurement and causality have been based upon those cited in Hair et al. (2016). For reflective measurement, measures represent the impact of related constructs. For PLS-SEM, causality is from the construct (depicted with arrows) to its measures (which are called indicators). These indicators can correspond to a survey question. Conversely, formative measurement indicators fully form or cause the construct, with arrows pointing from the indicators to the construct. Causality means different things in relation to differing methods and methodologies. When using PLS-SEM (Hair et al., 2016), causality is referred to when considering causal indicators. Causal indicators (used in formative measurement models) “cause” the latent variable. Hence, causal indicators need be in alignment with a theoretical definition of the concept under investigation. That is why a CR review was conducted in order to create the survey questions (that correspond to the indicators) and to predict the latent variables involved.

5.6.1 Stage1: Structural Model Specification

Stage one of the PLS-SEM approach is specific to the structural model and whittles the model down to pathways and variables that are significant. Stage one involved thinking through (via CR literature review) the latent variables and their measured indicators (survey questions). Figure 5-23 shows the original model built from the CR evaluation of the literature.

After running the data through SMART-PLS 3.0, it was found that enjoyment and engagement closely overlapped, and the enjoyment pathway was not found to be significant. Therefore, enjoyment was adjusted to become an indicator of the engagement latent variable and was later removed altogether due to lack of statistical significance.

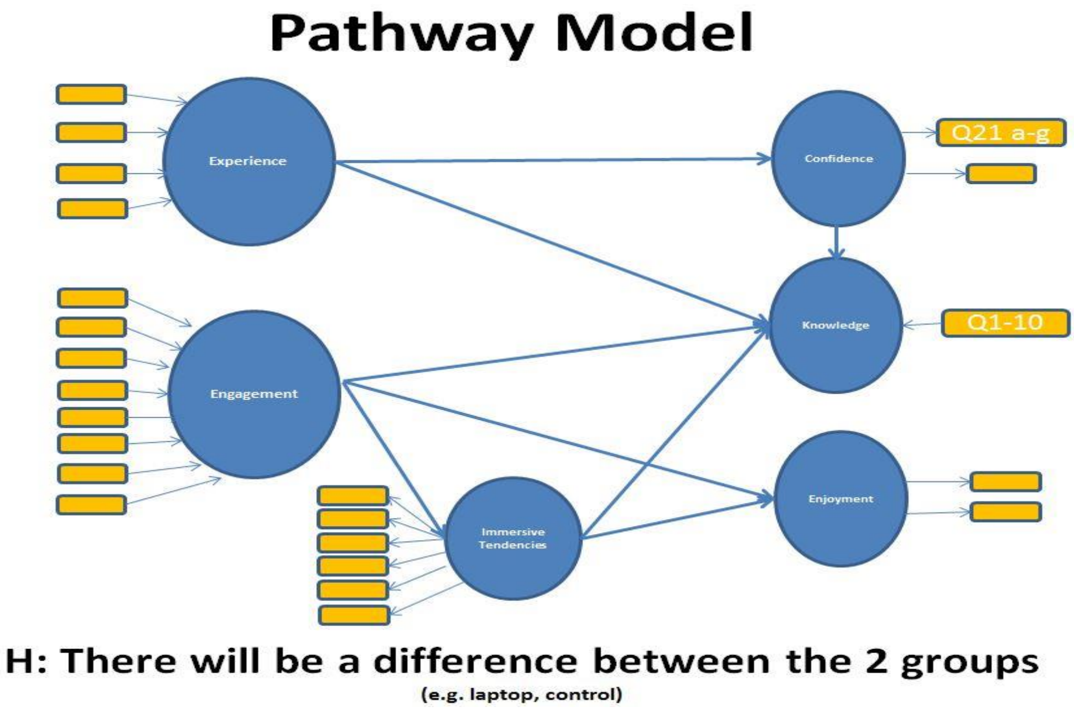


Figure 5-23 Original Proposed Conceptual Model which had Included Enjoyment as its own Construct

The final structural model (shown in Figure 5-24) contains three formative latent variables (engagement, experience and confidence), one reflective variable (immersion) and one single item dependent variable (knowledge). Knowledge has been included as a single item latent variable because it represents the summed totals of the hypoglycaemia MCQ scores (out of a total of 10 points), completed before and after the students used their deteriorating patient case study exercise.

5.6.2 Stage 2: Measurement Model Specification

The second stage of Hair et al.'s (2016) approach involved adding the indicators to the variables using the SMART-PLS 3.0 software. This meant connecting the regression pathways, which involved making formative and reflective decisions. The post-test survey data from the BOSs were exported into Excel and then SPSS. Once in SPSS, question headings were renamed as indicators and the data were formatted ready to be imported into SMART-PLS 3.0 software. Once in SMART-PLS 3.0, indicator data were converted into the pathway model via a simple drag and drop approach.

5.6.3 Stage 3: Data Collection and Screening

When using PLS-SEM, as with many statistical approaches, procedures must be followed in the case of any missing values in the data set. However, due to the choice of settings within the BOSs, there were no missing values on the final Excel data set. According to procedures, the data must also be examined (Hair et al., 2016) for any suspicious response patterns or outliers, none of which were located for the data set used in this study.

5.6.4 Stage 4: Model Estimation and the PLS-SEM Algorithm

Figure 5-24 shows the final pathway model with Path Coefficients and P-*values* on the Inner Model and Outer Weights/loadings and *P-values* on the Outer Model.

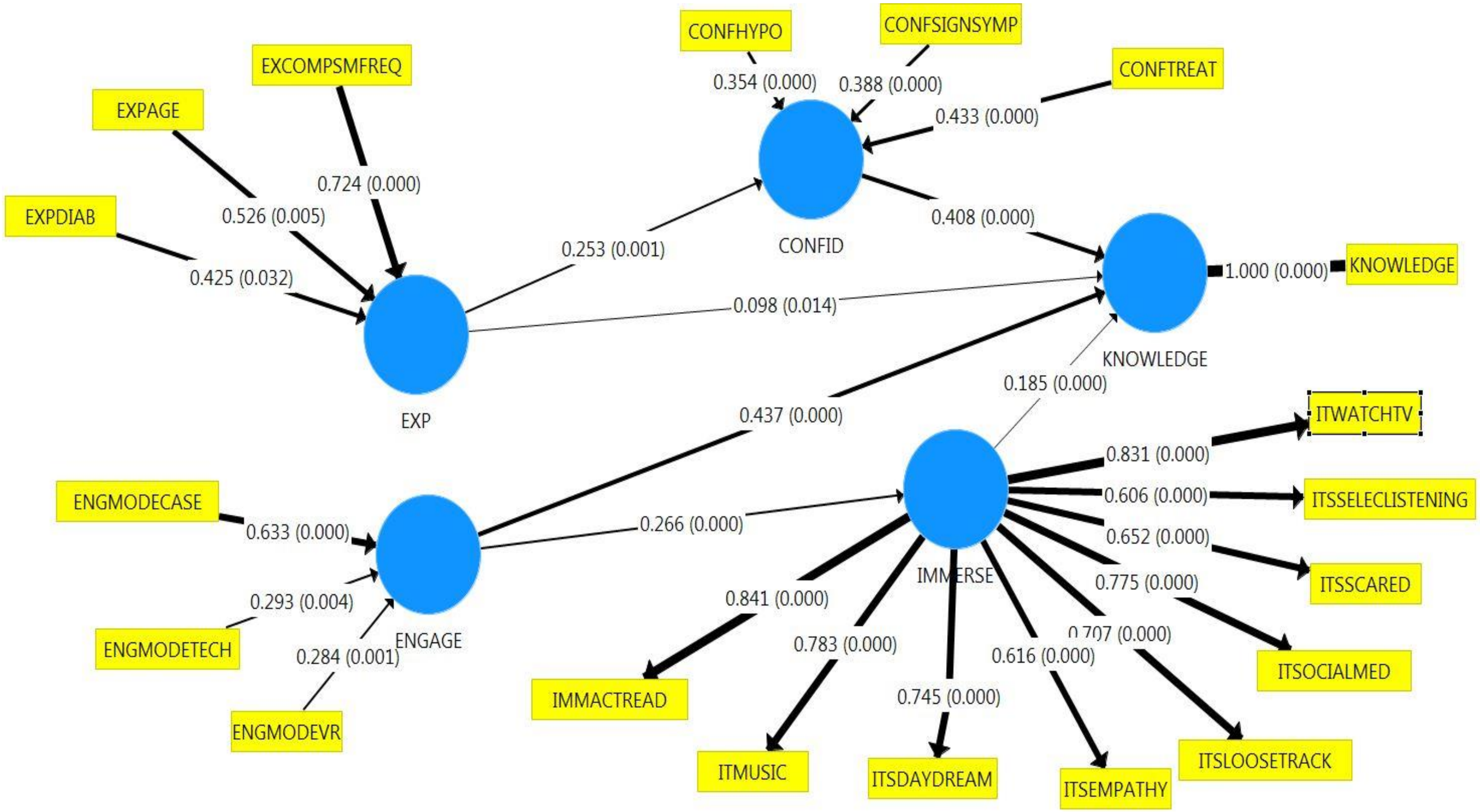


Figure 5-24 Conceptual Model

Figure 5-25 demonstrates how the Conceptual Model relates to the proposed CMPAO Configurations. The contextual mechanisms are the amount of diabetic nursing experience (that the students have) and relevant computing experience, and the participant engagement level with the learning content. The Programme mechanisms are the affordances of VR (including opportunities to practise via experiential learning and the ability to visualise abstract concepts) and the affordances of computing exercise, e.g. immediate feedback given and the chance to repeat exercise. The agents are confidence boosting via feedback and participant immersion with the desktop VR. Finally, the outcome being tested via quantitative analysis is improved knowledge of complex diabetic issues.

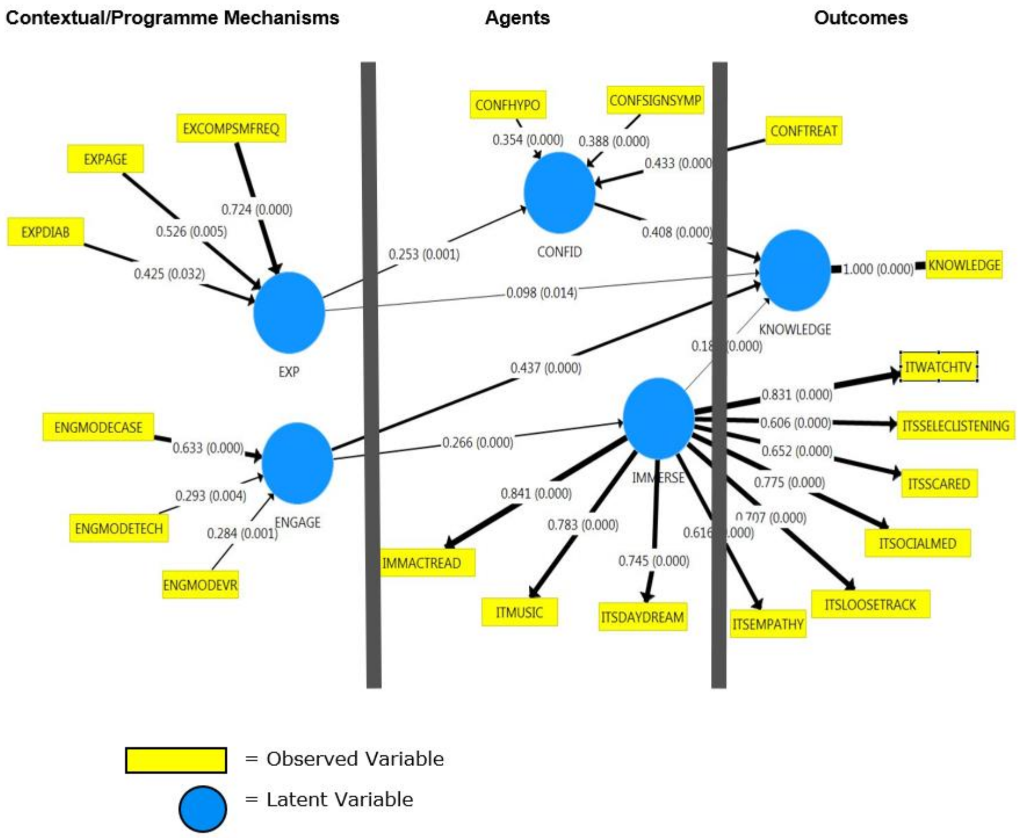


Figure 5-25 Overlay of Conceptual Model with CMPAO Configuration

Table 5.2 has been constructed to demonstrate how the indicators (yellow rectangles) related to the pre and post-test survey questions. The question number varies between the pre and post-test surveys because the consent questions were not repeated in the post-test. Survey questions that do not appear as indicators were removed because they were not found to be significant (Hair et al., 2016). Whilst, some questions that had been derived from the Immersive Tendencies Questionnaire (Witmer and Singer 1998) remained, all the TAM (Davis et al., 1989) survey questions were removed due to insignificance at this point.

Table 5-2 How indicators linked to pre and post-test surveys

Survey Question Number	Indicator	What the indicator refers to:	Previous survey from which the question was drawn
Q12	EXPAGE	Age of participant	
Q20	EXPDIAB	Participant experience of nursing diabetic patients	
Q15.1	EXPCOMPSMFREQ	How often participants use social media (a reflection of their digital/computing usage)	Falconer et al. (2018)
Q31.2	ENGAGEMODECASE	How engaging participants find learning from case studies	
Q31.3	ENGAGEMODETECH	How engaging participants find learning using interactive mobile device-based activities	
Q31.4	ENGAGEMODEVR	How engaging participants find using VR learning	
Q21.2	CONFHYPO	How confident participants feel about their knowledge of hypoglycaemia	
Q21.5	CONFSIGNSYMP	How confident participants feel to identify the signs and symptoms of hypoglycaemia	
Q21.7	CONFTREAT	How confident they feel to treat hypoglycaemia	
Q28.1	IMMACTREAD	How often the participants partake in reading	Witmer and Singer (1989)
Q29.6	ITMUSIC	Immersive tendency rating when listening to music	
Q30.4	ITSDAYDREAM	Liability to become immersed in a daydream	
Q30.6	ITSEMPATHY	Immersive tendency rating concerning empathy with characters in a book/movie	
Q30.1	ITSLOOSETRACK	Immersive tendency rating concerning losing track of time	
Q29.4	ITSOCIALMED	Immersive tendency rating concerning use of social media	
Q30.2	ITSCARED	Immersive tendency rating concerning remaining scared after watching a scary film	
Q30.7	ITSSELECLISTENING	Immersive tendency rating concerning selective listening	
Q29.2	ITWATCHTV	Immersive tendency rating concerning TV viewing	
Qs 2-11(For post-test Qs1-10)	KNOWLEDGE	10 questions relating to hypoglycaemia knowledge	

5.6.5 Stage 5: Systematic Evaluation of the Measurement Models

The first step in evaluating PLS-SEM results involves examining the quality of the measurement models and optimising indicator sets. According to Hair et al. (2016), the measurement model assessment includes:

Table 5-3 Reflective and Formative Measurement Model Assessment Processes

Variable Type	**Evaluative Tests**
Reflective	loadings, Cronbach's alpha/composite reliability/Ave, HTMT
Formative	redundancy analysis, VIF, and significance and relevance of the indicator weights

Reflective variable evaluation must be conducted prior to evaluation of the formative variables; hence Section 5.6 details the process undertaken to evaluate the reflective variables involved in the study.

According to Hair et al., (2016) there are three steps to evaluating reflective measurement models, namely: internal consistency, convergent validity (indicator reliability) and discriminant validity. The reflective variable included in the model was immersion, so this variable is assessed in the following sections.

The first stage in evaluating the reflective measurement models is examining consistency reliability. Internal consistency reliability is most frequently assessed using composite reliability (Werts et al., 1978). Upper values usually signpost upper levels of reliability. For an example, see Table 5-4.

Table 5-4 Internal Consistency Reliability Value Ranges

	Acceptable	**Satisfactory to Good**	**Problematic**
Reliability Values	0.60-0.70	0.70-0.90	>0.95

Item redundancy is indicated by values of higher than 0.95 and this reduces construct validity (Diamantopoulos et al., 2012). Reliability values of 0.95 and beyond also indicate the probability of disagreeable response configurations (e.g., straight lining), thereby triggering overstated correlations among the indicators' error terms (Hair et al., 2018). In the present study, evaluation of the composite reliability revealed that the immersion composite reliability value (0.907) exceeded the threshold and high level of internal consistency reliability. Cronbach's alpha is another measure of internal consistency reliability that assumes comparable thresholds but produces lower values than composite reliability (Hair et al., 2018). Construct reliability and validity revealed a Cronbach's alpha score for immersion of 0.878 and was hence over the 0.70 threshold.

The second stage in appraising the reflective measurement models is to assess the convergent validity. Convergent validity is the degree to which the construct joins to clarify the variance of its items (Hair et al., 2018). The metric used for assessing a construct's convergent validity is the average variance extracted (AVE) for all items on each construct. In order to calculate the AVE, the loading of each indicator on a construct must be squared, and then the mean value must be computed. An acceptable AVE is 0.50 or higher, indicating that the construct explains at least 50% of the variance of its items (Hair et al., 2018). The immersion variable was found to be above this criterion (0.619), thus, the measure of the reflective construct has pronounced levels of convergent validity.

Outer loadings are also examined as part of convergent validity. Loadings above 0.708 are suggested, as they indicate that the construct explains more than 50% of the indicator's variance, consequently providing satisfactory item reliability (Hair et al., 2018). There was one reflective latent variable included in the model. Most outer loadings of the reflective construct, immersion, were above the value of 0.70, and the remaining three loadings were close to the 0.70 value (Figure 5-25). This suggests sufficient levels of indicator reliability (Hair et al., 2016). The highest loadings were for: IMMACTREAD (0.841) and ITWATCHTV (0.831).

IMMACT...	0.841
ITMUSIC	0.783
ITSDAYD...	0.745
ITSEMPA...	0.616
ITSLOOS...	0.707
ITSOCIAL...	0.775
ITSSCARED	0.652
ITSSELEC...	0.606
ITWATCH...	0.831

Figure 5-26 Outer Loadings for the IMMERSION Reflective construct demonstrating sufficient levels of indicator reliability

Finally, the third step in evaluating the reflective measurement models is probing discriminant validity. Discriminant validity is the extent to which a construct is empirically distinct from other constructs in the structural model (Hair et al., 2018). It can be assessed using the Heterotrait-Monotrait (HTMT) ratio of the relationships (Voorhees et al., 2016). The HTMT is defined as the mean value of the item correlations across constructs comparative to the mean of the average correlations for the items quantifying the same construct (Hair et al., 2018). Discriminant validity problems exist when HTMT values are elevated. Hensley and Hubona (2016) proposed a threshold value of 0.90 for structural models with constructs that are theoretically very comparable. An HTMT value above 0.90 would indicate that discriminant validity is not present. But when constructs are conceptually more diverse, a lower, more conservative, threshold value is suggested, such as 0.85 (Hair et al., 2018). This conservative threshold of 0.85 was selected for the current study. The HTMT value for the reflective construct (0.389) was lower than the threshold of 0.85 and hence supports discriminant validity.

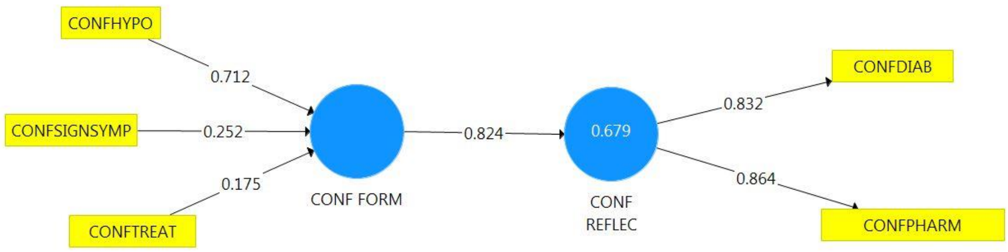
In summary, regarding the reflective measurement model assessment, all model assessment criteria were met, providing support for the measures' reliability and validity. Having evaluated the reflective variable (immersion), the procedure progressed to assessment of the formative variables.

The pathway model constructed in the present study contains four formative constructs (experience, engagement, confidence, knowledge). When formative constructs are incorporated in the structural model, PLS-SEM is the favoured method (Hair et al., 2018). Formative measurement models are gauged based on the ensuing: convergent validity, indicator collinearity, statistical significance, and significance of the indicator weights (Hair et al., 2016).To evaluate the formative measurement models, Hair et al.'s (2016) formative assessment procedures were followed and are described in the following sections.

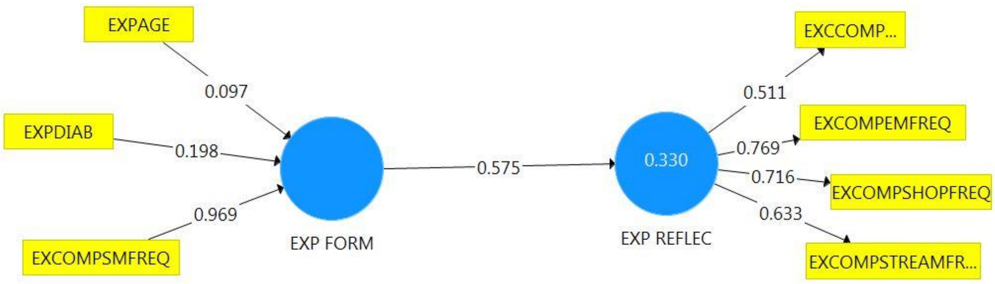
Convergent validity of the formatively measured constructs is tested by redundancy analysis, whereby the correlation of the construct is compared to a different measure of the equivalent concept (Tajik et al., 2020). In order to execute the redundancy analysis procedure, alternative reflectively measured indicators of the equivalent concept were included in the survey at the design stage. Cheah (2017) demonstrated that a single item, which is a good representative of the construct under deliberation, is generally adequate as a substitute measure. Hair et al. (2016) suggested that the correlation of the formatively measured construct with the single-item construct, measuring a similar concept, should be 0.70 or higher. Hence, in order to examine whether formative constructs exhibited convergent validity in the current study, separate redundancy analyses (Figure 5-26) were conducted for each construct (experience, confidence and engagement). As can be seen, this analysis yielded paths' coefficients that were higher than (e.g., confidence and engagement) or very close to (e.g., experience) the recommended threshold of 0.70, thus providing support for the formative constructs' convergent validity.

Redundancy Analysis Assessment of Formative Measurement Models

Convergent Validity of CONFIDENCE



Convergent Validity of EXPERIENCE



Convergent Validity of ENGAGEMENT

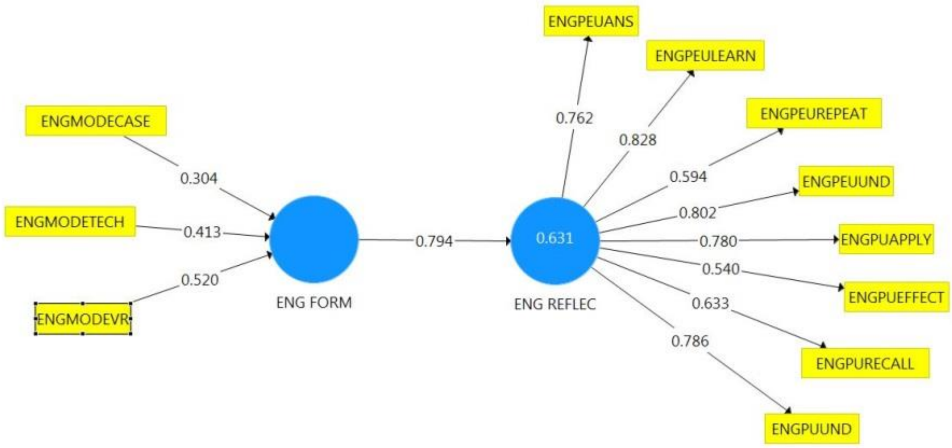


Figure 5-27 Convergent Validity of Formative Measurement Models- retained indicators shown on right hand side

The next step involved checking the formative measurement models for collinearity of indicators by examining the indicator variance inflation factor (VIF) values. VIF values were uniformly (1.0) below the threshold value of five. VIF values of five or beyond indicate acute collinearity issues amongst the indicators of formatively measured constructs. Nonetheless, collinearity concerns can also ensue at lesser VIF values of three (Thompson et al., 2017). Preferably, the VIF values should be near to three and lower. The conclusion was, therefore, that collinearity did not reach acute levels in any of the formative constructs and was not a concern for the approximation of the PLS-SEM pathway model.

In the third and final step, the indicator weights' statistical significance and relevance (i.e. size) were assessed. Bootstrapping is used to determine statistical significance because PLS-SEM is a nonparametric technique (Costa and Monteiro, 2018). Hair et al. (2016) suggested using bootstrap confidence intervals for significance testing in case the bootstrap distribution of the indicator weights is skewed. If the confidence interval of an indicator weight includes zero, this indicates that the weight is not statistically significant, and the indicator should be considered for removal from the measurement model (Hair et al., 2018). Outer weights were examined for their significance and relevance, using 5,000 bootstrapped examples. The p values in the formative measurement models displayed in Figure 5-27 were lower than 0.05 and therefore were established as significant outer weights at a significance level of 5% (i.e. $\alpha = 0.05$).

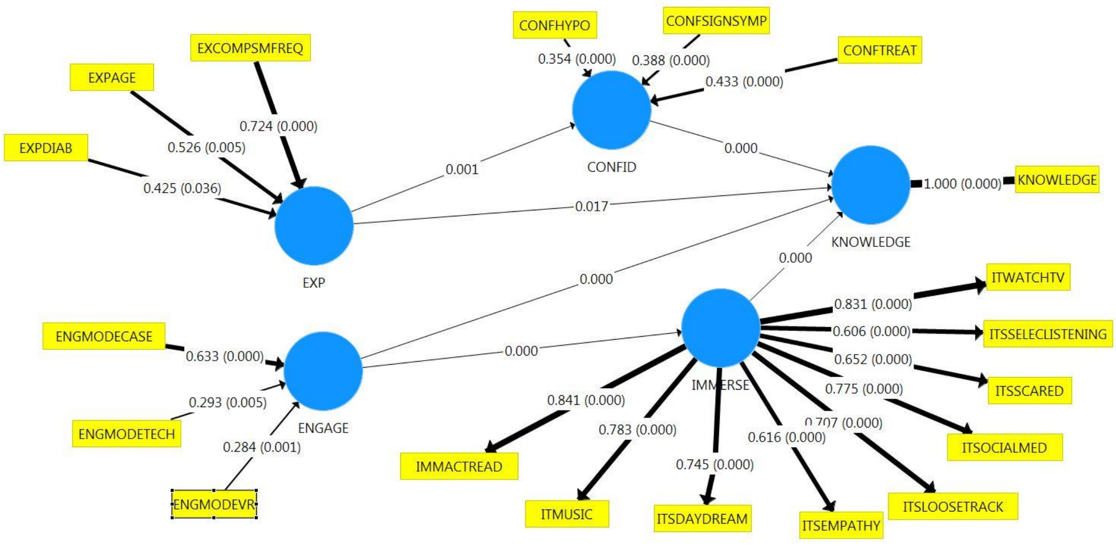


Figure 5-28 Bootstrapping *P* values and outer weights in the Modelling Window

Once the statistical significance of indicator weight assessment had taken place, each indicator was examined for relevance. Indicator weights values lie between -1 and +1 as standard, though they can fall outside this range (this indicates an abnormal result, e.g., small sample size of collinearity issues). Weights closer to 0 indicate weak relationships whilst those closer to +1 (or -1) suggest strong positive (or negative) relationships (Hair et al., 2018). As can be seen in Figure 5-27 the following indicators had the strongest positive relationships:

- IMMACTREAD (indicator weight= 0.841)
- ITWACHTV (indicator weight= 0.831)
- ITMUSIC (indicator weight= 0.783)
- ITSOCIALMED (indicator weight= 0.775)

According to Hair et al. (2016), indicators with a non-significant weight should certainly be removed if the loading is also not significant. A small but significant loading of 0.50 and below indicates that one ought to contemplate removing the indicator, except if there is robust backing for its presence on the grounds of measurement theory. Nonetheless, if an indicator weight is not significant, it is not automatically interpreted as evidence of reduced measurement model quality. In its place, the indicator's entire impact on the construct is deliberated (Hair et al., 2012), as defined by its outer loading. Outer loading is the bivariate correlation between the indicator and its construct (Hair et al., 2018). SMART-PLS software also calculates outer loadings. The *P* values of the outer loadings were below 0.01, suggesting that all loadings were significant at the level of 1% and the EXPDIAB indicator was significant at the 5% level. Thus, the indicators were retained in the formative constructs despite some indicators having insignificant weights.

Careful decisions were made regarding whether to remove formative indicators grounded on statistical outcomes. The first consideration was that the larger the quantity of indicators, the lesser their average weight. Therefore, formative measurement models are intrinsically restricted in the quantity of indicator weights that can be statistically significant (Hair et al., 2012). Secondly, formative measurement principles necessitate the indicators to entirely capture the complete realm of a construct, as demarcated by the investigator in the conceptualisation phase; hence they should rarely be removed. Unlike reflective measurement models, formative indicators are not substitutable and eliminating even a single indicator can thus diminish the measurement model's content validity (Diamantopoulos et al., 2012). The analysis of the outer weights and loadings concluded the assessment of the formative measurement models. The measurement models met all the required criteria; therefore, the evaluation proceeded to review the structural model (Hair et al., 2016).

5.6.6 Stage 6 Evaluation of the Structural Model

All reflective and formative constructs exhibited acceptable levels of quality. Thus, evaluation of the structural model was conducted, using the following ways: collinearity assessment, coefficients of determination (R^2) (and f^2 effect sizes), predictive relevance (Q^2), and the statistical significance and relevance of path coefficients (Hair et al., 2016). This stage reports the evaluations and assessments conducted in order to demonstrate the validity and reliability of the structural model. Stage 6 also included computation of the configural structural model and deletion of non-significant paths.

Structural model path coefficients are derived from approximating a sequence of regression equations. Prior to this structural relationship assessment collinearity needs to be scrutinised to ensure it does not bias the regression outcomes. This procedure is comparable to evaluating formative measurement models, but the latent variable scores of the predictor constructs in a partial regression are used to calculate the VIF values (Hair et al., 2018). Collinearity issues among predictor constructs are indicated if VIF values are above five. For the present study, VIF values were examined for all sets of predictor constructs in the structural model.

Figure 5-28 shows the VIF values of all amalgamations of endogenous constructs (represented by the columns) and equivalent exogenous (i.e. predictor) constructs (represented by the rows).

	CONFID	ENGAGE	EXP_	IMMERSE	KNOWLE...
CONFID					1.935
ENGAGE				1.000	2.023
EXP_	1.000				1.110
IMMERSE					1.097
KNOWLE...					

Figure 5-29 VIF Values in the Structural Model showing all values below the threshold of five

As can be seen, all VIF values were below the threshold of five. Therefore, collinearity among the predictor constructs was not a critical issue in the structural model, and examination of the report results could continue.

Having established that collinearity was not a concern; the next phase is investigating the R^2 value of the endogenous construct(s). The R^2 measures the variance, which is explained in each of the endogenous constructs and is consequently a measure of the model's explanatory power (Shmueli and Koppius, 2011). Rigdon (2012) referred to the R^2 as in-sample predictive power. The R^2 ranges from 0 to 1, with higher values signifying a superior explanatory power. Henseler and Hubona (2016) provided the following guidelines: R^2 values of 0.75, 0.50 and 0.25 can be considered substantial, moderate and weak, respectively. The R^2 is a function of the number of predictor constructs – the larger the number of predictor constructs, the higher the R^2. For the current study, the R^2 value for knowledge, i.e. 0.767, was considered moderate, whereas immersion and confidence at 0.071 and 0.064 respectively were weak.

The next procedure was to evaluate how the deletion of a predictor construct affected an endogenous construct's R^2 value. This metric was the f^2 effect size and is redundant to the extent of the path coefficients. More precisely, the rank order of the predictor constructs' significance in explaining a dependent construct in the structural model is often the same when equating the size of the path coefficients and the f^2 effect sizes. Nitzl (2016) stated that if the rank order of the constructs' significance, when clarifying a dependent construct in the structural model, varies when equating the size of the path coefficients and the f^2 effect sizes, then the f^2 effect size must be recounted to explain the manifestation of, for example, partial or full mediation. Usually, values higher than 0.02, 0.15 and 0.35 depict small, medium and large f^2 effect sizes, respectively (Cohen et al., 2017). Figure 5-29 shows the f^2 values for all combinations of constructs in the current study. Confidence and engagement had a large effect on knowledge, whilst immersion had a medium effect and experience had the smallest effect.

	CONFID	ENGAGE	EXP_	IMMERSE	KNOWLE...
CONFID					**0.371**
ENGAGE				0.076	**0.405**
EXP_	0.068				0.037
IMMERSE					0.134
KNOWLE...					

Figure 5-30 f^2 Values for all combinations of constructs in the current study, showing that Confidence and Engagement had a Large Effect on Knowledge

Using Total Effects, it was possible to evaluate how strongly each of the driver constructs (experience and engagement) influenced the key target variable knowledge via the mediator constructs of confidence and immersion. The total effects are shown in Figure 5-30. Each column represents a target construct, whereas the rows represent predecessor constructs. For example, regarding knowledge, we can see that engagement had the strongest total effect on knowledge (0.486), followed by confidence (0.408) and experience (0.201). Once again, this might point to only a partial mediation of engagement via immersion.

	CONFID	ENGAGE	EXP_	IMMERSE	KNOWLE...
CONFID					0.408
ENGAGE				0.266	0.486
EXP_	0.253				0.201
IMMERSE					0.185
KNOWLE...					

Figure 5-31 Total Effects, which Indicated only a Partial Mediation of Engagement via Immersion

The Q^2 metric is based on a blindfolding process, used for predictive relevance evaluation, that eliminates single points in the data matrix, attributes the removed points with the mean and approximates the model parameters (Rigdon, 2012). As such, the Q^2 combines aspects of out-of-sample prediction and in-sample explanatory power (Sarstedt and Cheah, 2019). Using these approximations as input, the blindfolding process predicts the data points that were eliminated for all variables. Smaller differences between the predicted and original values lead to a higher Q^2 value, hence indicating a higher predictive accuracy. Generally, Q^2 values should be greater than zero for an exact endogenous construct to indicate the predictive accuracy of the structural model for that construct (Hair et al., 2019). Q^2 values greater than 0, 0.25 and 0.50 depict small, medium and large predictive relevance of the PLS-path model, respectively. Analogous to the f^2 effect sizes, it is possible to compute and interpret the q^2 effect sizes. For the present study, as can be seen in Figure 5-31, knowledge has the highest q^2 value (0.721), followed by confidence (0.033) and immersion (0.028). Hence knowledge has large predictive relevance, whilst the other two latent variables shown have small predictive relevance. This might be as would have been expected because the knowledge indicator is a combination of the answers to ten questions.

Construct Crossvalidated Redundancy

Total | Case1 | Case2 | Case3

	SSO	SSE	Q^2 (=1-S...
CONFID	513.000	495.823	0.033
ENGAGE	513.000	513.000	
EXP_	513.000	513.000	
IMMERSE	1,539.000	1,495.174	0.028
KNOWLE...	171.000	47.666	0.721

Figure 5-32 Q^2 Values showing Knowledge as having the Largest Predictive Relevance

Having confirmed the model's explanatory power and predictive power, the ultimate step involved assessing the statistical significance and relevance of the path coefficients. The analysis of the path coefficients is equivalent to that of the formative indicator weights (Hair et al., 2018). Therefore, bootstrapping was run to evaluate the path coefficients' significance and assess the values (typically falling in the range of −1 and +1). The construct's indirect effects on a certain target construct, via one or more intervening constructs, were interpreted (Nitzl, 2016). By taking the construct's indicator weights into consideration, it was possible to identify which specific element of engagement had had the strongest impact. Looking at the outer weights revealed that ENGMODECASE has the highest outer weight (0.633). This related to the survey item "I would engage well with learning from a case study" which involved participants rating their answer using a Likert scale. The examination of structural model relationships indicated that several path coefficients (e.g., experience → knowledge) had rather low values. This might be because it was mediated by confidence. Figure 5-32 shows the *P* values for the structural model relationships.

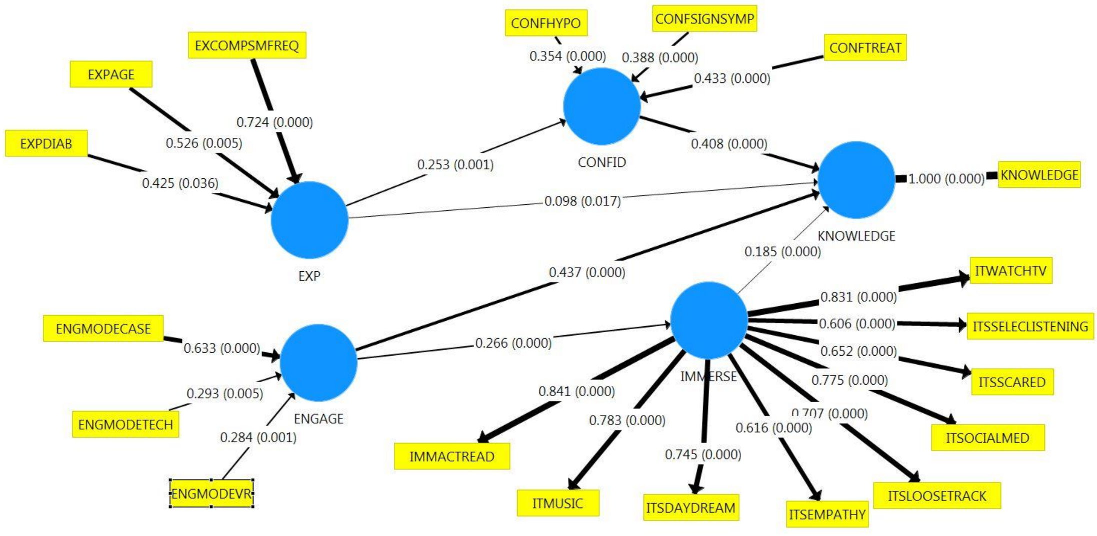


Figure 5-33 Bootstrapping Results showing Structural Model P values in the Modelling Window.

Assuming a 5% significance level, the findings were that all the relationships in the structural model were significant. Figure 5-32 shows partial mediation by both immersion and confidence. Finally, the results of the HTMT ratio for knowledge (0.492 at the 95% confidence level) confirm that values of the HTMT were below the threshold of 0.85 or 0.90 (the procedure did not produce results for other latent variables, Henseler et al. (2019). This means that the indicators for knowledge were distinct from other constructs in the structural model and hence discriminant validity of the knowledge indicator was confirmed.

To summarise, the above evaluations and assessments demonstrated the validity and reliability of the structural model. This means that interpretations taken from the results of the PLS-SEM model can be deemed to be accurate. Therefore, the PLS-SEM procedure could progress to the next stage which involved running through the data to compare the Control and Experimental Groups.

5.6.7 Invariance Assessment

The configural model was fitted to the pooled data. Having established that the PLS-SEM structural and measurement models are above the thresholds for validity and reliability, according to Hair et al. (2016), the next process involved group comparison. However, measurement invariance must be ensured (Hair et al., 2018), prior to equating group-specific parameter estimates for significant differences using a multigroup investigation. By establishing measurement invariance, it is evidenced that group differences in model estimates do not result from distinguishing content/meaning of the latent variables across groups. According to Horn and McArdle (1992) measurement invariance refers to:

> *“whether or not, under different conditions of observing and studying phenomena, measurement operations yield measures of the same attribute.”*

Failure to establish data equivalence can lead to measurement error, which would reduce the power of statistical tests of the hypotheses and also provide misleading findings and conclusions (Hult and Ketchen Jr, 2001). Therefore, an evaluation was conducted, using the MICOM to ascertain whether the effects in the model differed significantly for participants who were in the Control or Experimental Groups. This invariance analysis follows Henseler and Hubona's (2016) procedure for measurement invariance assessment in composite modelling, which contrasts group-specific measurement model estimates with those obtained from a model estimation using pooled data. Implementing Henseler and Hubona's (2016) endorsements, a three-step MICOM procedure was followed including i) configural invariance, ii) compositional invariance, and iii) scalar invariance (equality of composite means and variances). The three-step process was carried out using SMART-PLS 3. Table 5-5 provides a summary of the MICOM results.

Table 5-5 Summary of the MICOM via PLS-SEM Results Demonstrating Configural and Compositional Invariances

<table>
<tr><td colspan="5">Summary of the MICOM via PLS-SEM Results</td></tr>
<tr><td colspan="5">MICOM Step 1- Configural invariance</td></tr>
<tr><td colspan="5">Configural invariance established? YES</td></tr>
<tr><td colspan="5">MICOM Step 2- Compositional invariance established? YES</td></tr>
<tr><td>Composite</td><td>Correlation c</td><td>5% quartile of the empirical distribution of c_μ</td><td>P value</td><td>Compositional invariance established?</td></tr>
<tr><td>CONFIDENCE</td><td>0.993</td><td>0.842</td><td>.876</td><td>Yes</td></tr>
<tr><td>ENGAGE</td><td>0.969</td><td>0.874</td><td>.497</td><td>Yes</td></tr>
<tr><td>EXPERIENCE</td><td>0.975</td><td>0.506</td><td>.860</td><td>Yes</td></tr>
<tr><td>IMMERSE</td><td>0.994</td><td>0.968</td><td>.842</td><td>Yes</td></tr>
<tr><td>KNOWLEDGE</td><td colspan="4">Single Item Measure- hence outer relationship is 1 by design.</td></tr>
<tr><td colspan="5">MICOM Step 3- Assessment of Equality of Composite Means and Variances</td></tr>
<tr><td></td><td colspan="2">Difference of the composite’s mean value (=0)</td><td>Permutation
P value</td><td>Equal mean values?</td></tr>
<tr><td>CONFIDENCE</td><td colspan="2">1.347</td><td><.001</td><td>No</td></tr>
<tr><td>ENGAGE</td><td colspan="2">1.224</td><td><.001</td><td>No</td></tr>
<tr><td>EXPERIENCE</td><td colspan="2">0.268</td><td>.062</td><td>Yes</td></tr>
<tr><td>IMMERSE</td><td colspan="2">-0.455</td><td>.003</td><td>No</td></tr>
<tr><td>KNOWLEDGE</td><td colspan="2">1.166</td><td><.001</td><td>No</td></tr>
<tr><td></td><td colspan="2">Logarithm of the composite’s variances ratio (=0)</td><td>Permutation
P value</td><td>Equal variances?</td></tr>
<tr><td>CONFIDENCE</td><td colspan="2">-0.694</td><td><.001</td><td>No</td></tr>
<tr><td>ENGAGE</td><td colspan="2">0.034</td><td>.827</td><td>Yes</td></tr>
<tr><td>EXPERIENCE</td><td colspan="2">0.247</td><td>.367</td><td>Yes</td></tr>
<tr><td>IMMERSE</td><td colspan="2">0.003</td><td>.992</td><td>Yes</td></tr>
<tr><td>KNOWLEDGE</td><td colspan="2">0.248</td><td>.179</td><td>Yes</td></tr>
</table>

As can be seen in Table 5-5 both Configural and Compositional invariances were established. MICOM Step 2, demonstrated measurement model invariance between groups; in other words, both groups tackled the survey in the same way conceptually (Henseler and Hubona, 2016).

5.6.8 Multi Group Analysis

Having demonstrated that MICOM Steps 1 and 2 were adhered to, multigroup invariance was examined. The third and final step of the MICOM comprises detecting the scalar invariance among the groups (Sinkovics et al., 2016). Looking at the results of the MICOM Step 3, the mean values hold the most importance. The constructs engagement, confidence and knowledge do not have equal mean values (scalar invariance was found), whereas, experience does have equal mean values, and a permutation p value of greater than 0.05. This means that, across the Control and Experimental Groups' participants', prior experience was consistent between groups (as we would expect). However, having completed the intervention the Experimental Group had higher mean values for engagement, confidence, and knowledge. A recap of the specific and testable hypotheses is shown in Table 5-6.

Table 5-6 Recap of the Specific Hypothesis

Hypothesis Number	***Specific Hypothesis***
H6	Path CONFIDENCE→KNOWLEDGE, $H_\circ$ $\beta=0$; (H_1 $\beta \neq 0$)
H7	Path *ENGAGE→IMMERSION*, $H_\circ$ $\beta=0$; (H_1 $\beta \neq 0$)
H8	Path *ENGAGE→KNOWLEDGE*, $H_\circ$ $\beta=0$; (H_1 $\beta \neq 0$)
H9	Path *EXP→CONFID*, $H_\circ$ $\beta=0$; (H_1 $\beta \neq 0$)
H10	Path *EXP→KNOWLEDGE*, $H_\circ$ $\beta=0$; (H_1 $\beta \neq 0$)
H11	*Path IMMERSE→KNOWLEDGE*, $H_\circ$ $\beta=0$; (H_1 $\beta \neq 0$)
H12	LV KNOWLEDGE, $H_\circ$ $Gp1_{mean} - Gp2_{mean} = 0$; ($H_1$ $Gp1_{mean} - Gp2_{mean} < 0$)
H13	LV CONFIDENCE, $H_\circ$ $Gp1_{mean} - Gp2_{mean} = 0$; ($H_1$ $Gp1_{mean} - Gp2_{mean} < 0$)
H14	LV ENGAGEMENT, $H_\circ$ $Gp1_{mean} - Gp2_{mean} = 0$; ($H_1$ $Gp1_{mean} - Gp2_{mean} < 0$)
H15	LV IMMERSION, $H_\circ$ $Gp1_{mean} - Gp2_{mean} = 0$; ($H_1$ $Gp1_{mean} - Gp2_{mean} < 0$)

The permutation test results (5,000 permutations) demonstrate that the mean value and variance of a composite in the Experimental Group do significantly diverge from the results in the Control Group. The permutation test reliably controls for Type 1 errors when the assignment of observations occurs randomly, as is the case when using SMART-PLS 3.0. A two-tailed 95% permutation-based confidence interval was created. According to Hair et al., (2018:154) if the original difference (d) of the group-specific path coefficient estimates does not fall into the confidence interval, it is statistically significant. Path coefficient null hypotheses for H6, H7, H8, H10 and H11 can all be accepted. H9 concerning the ENGAGE→IMMERSION has an original difference of 0.278; as this does not fall between the confidence intervals of -0.255 and 0.262, it is statistically significant. Hence, the null hypothesis for H9 was rejected. Table 5-7 shows the rejected null hypotheses as unshaded rows.

Table 5-7 Rejected Null Hypotheses

Hypothesis	***Path***	***Confidence Intervals***	***Original Difference***	***H_o Remark***
H6	*EXP-CONFID*	*-0.296, 0.294*	*-0.053*	*H_o accepted*
H7	*CONFID-KNOWLEDGE*	*-0.195, 0.197*	*-0.124*	*H_o accepted*
H8	*EXP-KNOWLEDGE*	*-0.156, 0.159*	*-0.048*	*H_o accepted*
H9	*ENGAGE-IMMERSE*	*-0.255, 0.262*	*0.278*	*H_o rejected*
H10	*IMMERSE-KNOWLEDGE*	*-0.149, 0.146*	*0.023*	*H_o accepted*
H11	*ENGAGE-KNOWLEDGE*	*-0.195, 0.195*	*0.005*	*H_o accepted*
H12	LV KNOWLEDGE, H_o $Gp1_{mean} - Gp2_{mean} = 0$; ($H_1$ $Gp1_{mean} - Gp2_{mean} < 0$)			*H_o rejected*
H13	LV CONFIDENCE, H_o $Gp1_{mean} - Gp2_{mean} = 0$; ($H_1$ $Gp1_{mean} - Gp2_{mean} < 0$)			*H_o rejected*
H14	LV ENGAGEMENT, H_o $Gp1_{mean} - Gp2_{mean} = 0$; ($H_1$ $Gp1_{mean} - Gp2_{mean} < 0$)			*H_o rejected*
H15	LV IMMERSION, H_o $Gp1_{mean} - Gp2_{mean} = 0$; ($H_1$ $Gp1_{mean} - Gp2_{mean} < 0$)			*H_o rejected*

Table 5-8 Key Finding

The Key Finding

is that the conceptual model shown below was proven to have ALL significant pathways across the data set (i.e. across the Control and Experimental Groups).

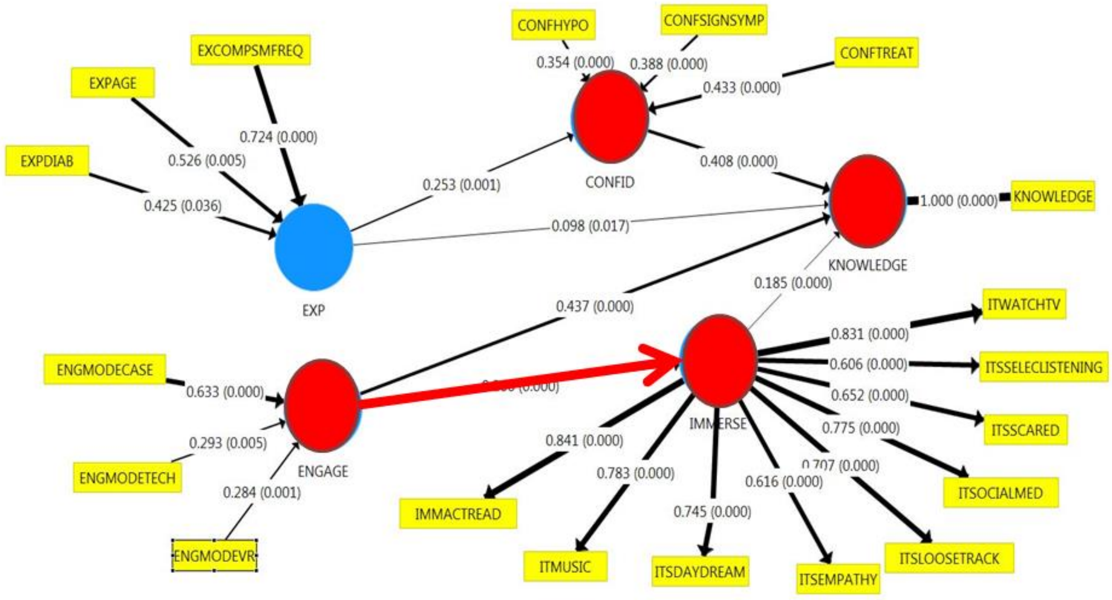


Figure 5-34 Group differences shown on the Conceptual Model- Between groups (Experimental versus Control) the significant differences are shown in red.

Across the data set (post-test survey) all pathways were found to be significant. This means that the experience (including prior computing and diabetic nursing experience) to knowledge pathway was partially mediated by confidence. This further means that for all the participants involved in the study, the more previous experience they had, the more confidence they had when using either diabetic intervention (paper or VR) and thus the higher their knowledge scores post intervention. The upper section of the diagram also shows (in red) that, across the groups, confidence and knowledge scores were higher in the Experimental Group post intervention.

In the lower half of the diagram, across the data set the engagement to knowledge pathway was partially mediated by immersion, and again these were significant pathways across the whole data set. Across groups however, the only significant pathway in the conceptual model was the engagement to knowledge pathway (shown as a red arrow). This may indicate that the improved knowledge scores for the Experimental Group have resulted from this “key pathway”, which could be deemed to be the “action point” of the PLS-SEM model. In other words, for the Experimental Group, whilst learning about diabetes using the VR simulation, students became more engaged in their learning and therefore more immersed in their learning than those learners who did not use VR (the Control Group); and this resulted in those students attaining higher knowledge scores (as measured by a set of ten MCQs) than the Control Group. This supports the previous research conducted, e.g., Tcha-Tokey et al. (2018) who suggested that engagement is the first step towards immersion.

The latent variable prior experience (shown in blue) was not found to statistically significantly different between groups. This means that when using the VR simulation, a student’s prior experience, including if they had used VR before, and so forth, did not impact on their MCQ scores. This indicates that the VR simulation is an inclusive learning tool, regardless of students’ age, computing experience or diabetic nursing experience. It means that it is a suitable learning method for all students (provided they are not susceptible to nausea or migraines and so forth). This finding is supported by the earlier work of Childs (2010) and Falconer et al. (2018) who concluded that prior experience does not affect the ability of a student to learn and have positive experiences when using VR.

5.6.9 The Main Outcomes of Model and Group Analysis

In summary, the results indicated that there was a difference between the Control and Experimental Groups across the PLS-SEM model. Across the groups’ participants’, prior experience was consistent between groups (as we would expect). However, a striking result is that having completed the case study the Experimental Group had higher mean values for engagement, confidence and knowledge. This means that the VR learning activity was, possibly though affordances of immediate feedback and visualisation and so forth, more engaging than normative instruction methods, made students feel more confident about their knowledge and improved their short-term learning outcomes in terms of hypoglycaemia quiz scores.

Not only did the PLS-SEM analysis evidence superior Experimental Group learning gain, confidence scores, and engagement levels, it was possible to pinpoint the precise point of action of the conceptual framework. Any pathway found to be significantly different between the Control and Experimental Groups can be viewed as being the “point of action” of the conceptual pathway model. The point of action for the pathway model appears to be the engagement to immersion pathway, which was found to be statistically significant (as measured by the permutation test). This means that via engaging and immersing students in their learning (e.g., via affordances of VR learning including experiential and visualisation learning opportunities), despite the software being low cost and low fidelity, better learning outcomes were obtained (Figure 5-34).

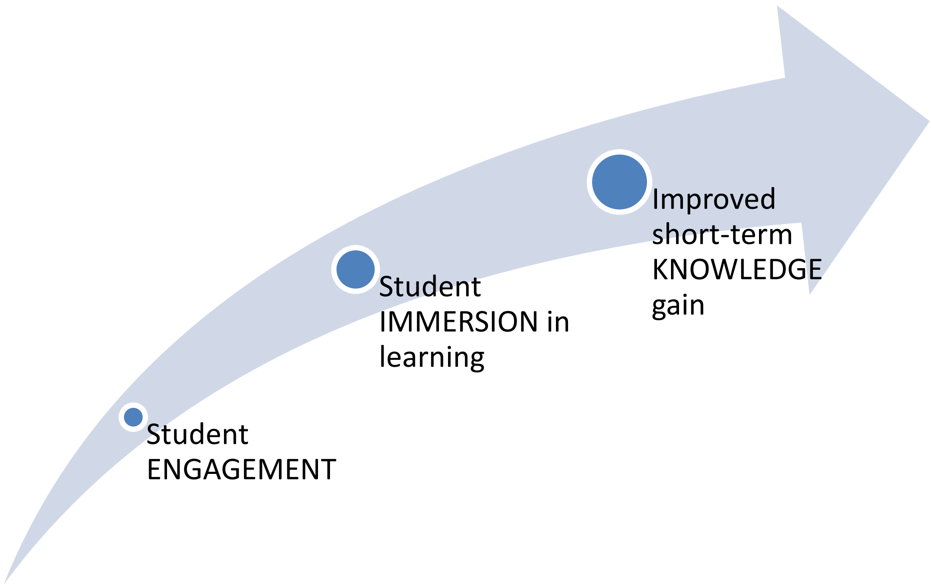


Figure 5-35 How VR Diabetes Innovation acted, as evidence by PLS-SEM analysis

Together these results provide important insights into how the use of VR learning compares to traditional or normative HE teaching methods and this will be explored in Chapter Seven.

5.7 Chapter Summary

This chapter presented the validity and reliability of the PLS-SEM pathway model at length, prior to running the pre and post-test survey data through the model. Having established configural and compositional invariance, an assessment of equality of composite means and variances was conducted. The results in this chapter indicate that not only was the VR intervention more enjoyable, moreover, the data suggest that it was more effective for student learning than the traditional learning methods. The next chapter, therefore, moves on to discuss the qualitative results. The Discussion Chapter (Seven) will synthesise both the quantitative and qualitative findings via a CR angle.

Chapter 6 Reporting of Qualitative Results

6.1 Synopsis

Having reported the quantitative results in Chapter Five, this chapter reports the qualitative findings, including FGD data, and open-ended survey question results. These qualitative results will be reported synergistically, under key themes derived from the CR review of the literature. Participant views concerning the advantages, challenges, and future uses of VR learning will also be presented. This chapter begins by reporting the demographics of the FGD findings.

6.2 Focus Groups' Demographics

In total, 21 participants took part in an FGD. These included:

- five LTs, who had been involved in supporting the students to complete the VR simulation during the four different lecture sessions.
- 15 nursing students (9% of the study population) also took part in FGDs.
- one mental health lecturer completed an individual interview.

The participant demographics are illustrated in Figure 6-1. The academic who had assisted with writing the diabetic case study, felt she had been too involved to be interviewed as her thoughts would be biased.

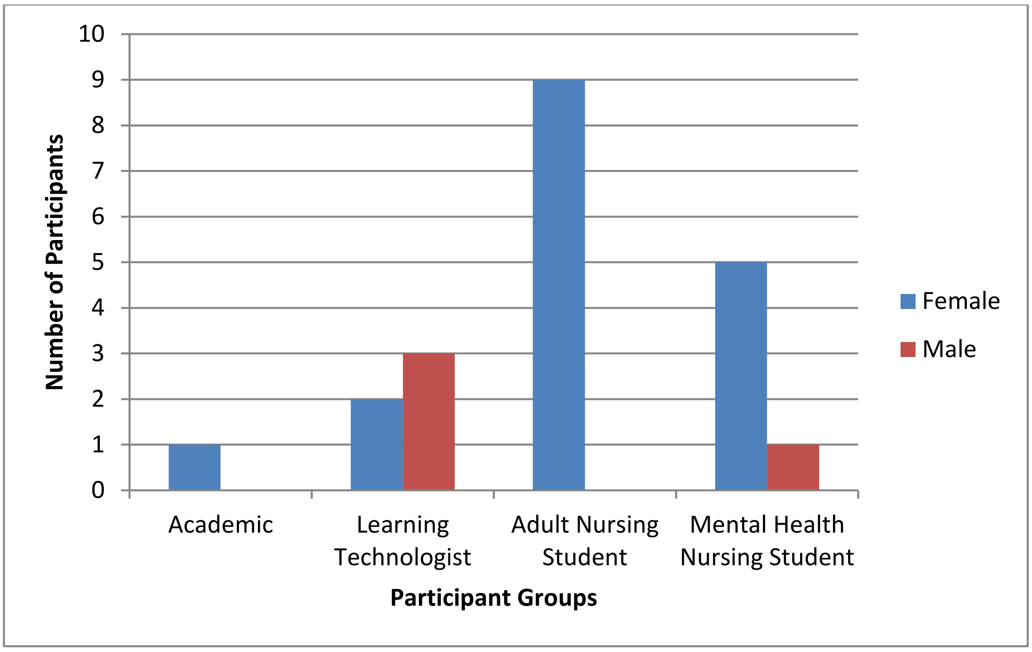


Figure 6-1 Demographics of the Focus Group Discussion

The 21 FGD/interview participants were asked the following three semi-structured questions:

Question 1. What was your experience of using the laptop case study?

Question 2. What were the challenges you faced, if any?

Question 3. How do you think it could be used in the future?

These questions overlap some of the survey questions and were designed to elicit more elaborate responses from the respondents.

6.3 Participant Views

Within this section the open-ended survey (n=171) and FGD responses from students (n=15), have been combined with FGD responses from LTs (n=5) and the academic (n=1) who had agreed to be interviewed. Responses about the advantages of the VR exercise have been categorised using the four CR themes (engagement, confidence, knowledge and immersion). Engagement based comments fell within two subcategories, namely "enjoyment" and "ease of use". The fifth CR theme (experience) did not surface in the survey or FGD comments, but this would be expected because respondents were not asked open questions about their prior experience (neither concerning computing nor nursing).

FGD respondents were not directly asked to share what they felt the advantages of the VR simulation were, but were asked the more general question: *"What was your experience of using the laptop case study?"* Table 6-1 compares responses from the academic, LTs and students who took part in the FGDs. The headings (e.g., engagement), are the categories that emerged from the review.

Table 6-1 A comparison of the response category frequency between respondent groups. Overall, immersion was the most frequent theme.

Participant	**Engagement**		**Confidence**	**Knowledge**	**Immersion**
	Enjoyment	**Ease of Use**			**(Bringing it to life)**
Academic	1	1	1	1	1
LT	5	5	1	4	5
Student-FGD	15	13	13	10	14
Student-Survey	4	7	7	4	20
Total	25	26	22	19	40

The most striking observation from the data is that comments classified using the engagement theme (51 references) were the most frequent. Within this theme, the most cited advantage of the simulation was "ease of use", and these answers included comments about being able to repeat the exercise. Comments relating to the immersion theme were the second most prevalent (n=40). Overall, each of the four CR themes were represented frequently, this suggests that they were valid choices (as suggested by the CR review).

Equally, there were no comments that were not reasonably put into any one of the four CR categories (other comments have been classified as referring to challenges or future uses of the VR exercise). When collated, comments referencing each theme were as shown in Figure 6-2.

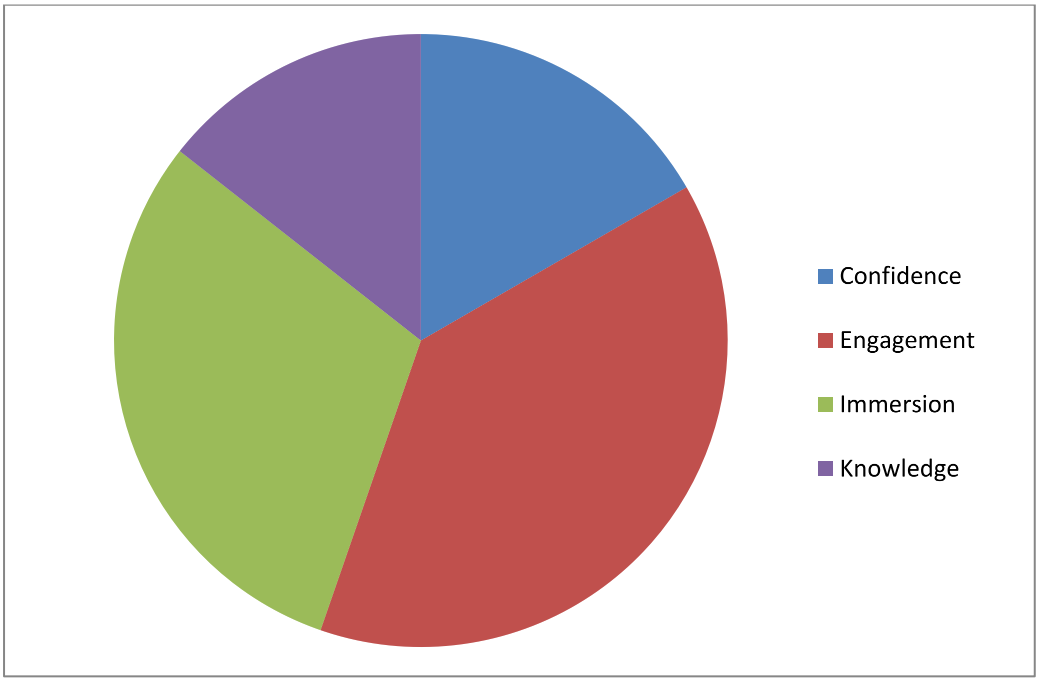


Figure 6-2 Frequency of each theme drawn out from responses from open-ended survey questions and FGDs

Focusing on the optional open-ended survey questions, the Experimental Group participants were asked the following open-ended question: *"What do you think the advantages (if any) of the laptop-based exercise are?"* All the responses to the open-ended question about the possible advantages of the simulation have been tabulated in Table 6-2.

Table 6-2 Advantages of VR simulation as perceived by survey participants

Advantage	**Frequency of survey participants citing this potential advantage**
Immersion	16
Instant feedback	5
Experiential learning	4
Aids memory	4
Interactive	4
Engaging	4
Can be repeated	3
Safe conditions in which to practise	2
Complements learning	1

The theme of immersion was identified in responses by 16 participants, with typical responses citing the "*real life*" and "*visual*" qualities of the simulation. The next most popular type of comment referred to the "*instant feedback*" that the computer software provided. Other frequent (n=4) comments related to: "*experiential learning*", "*aide-memoire*", the "*interactivity*" of the software, and finally that the laptop-based exercise was an engaging way to learn.

By way of comparison, participants who were randomised into the Control Group were asked a similar question: *"What do you think the advantages (if any) of the paper-based exercise are?"* Two students felt that there were no advantages. Four students responded that the paper exercise was quicker to complete in their opinion. Finally, six felt that the paper version was easy to use, understand and access: "*There is little that can go wrong with them*".

FGD participants were asked about what challenges the computer exercise might bring. The academic chose to discuss financial considerations in her comment. The comment below illustrates her previous experience with VR in an HE setting:

> *"I think it is about investment, so if you are going to do this work in an HE setting you have got to have investment, not just in terms of money but also in time and appreciating the work people do."*

The potential challenges identified by FGD LTs, FGD students and Experimental Group survey respondents have been combined in Table 6-3.

Table 6-3 Challenges of the VR simulation as expressed by participants

Challenges	**Frequency of LTs citing this challenge**	**Frequency of FGD students citing this challenge**	**Frequency of survey respondents citing this challenge**	**Total citations of the challenge**
Unreliable software/hardware	1	3	9	13
Navigating the avatar around the virtual space	1	2	5	8
Experience of computing needed	1		4	5
Students being reluctant to download apps	4			4
Downloading issues	2		3	5
Total citations of challenges made	9	5	21	NA

The data suggest that software/hardware concerns (not separated because the respondents did not discuss these separately) were raised by both LTs and students and were the main concern about the VR simulation. For example, students and LTs were worried that the software could make the computer crash. The ability or ease of use to move the avatar around the virtual wardroom was the second most prevalent concern amongst LTs and students. To illustrate this point, the avatar could become temporarily stuck behind the virtual television if the avatar moved behind where the tv was placed. These data imply that LTs (nine challenges cited by five LTs) were more likely to cite possible challenges than students were (26 challenges cited by 83 students). This might be because LTs need to find solutions to these challenges and are aware of the practical implications of using such software in large-scale situations.

Again, by way of comparison, the Control Group (via open-ended survey questions), were also asked if they had experienced challenges when using the paper-based exercise. Thirteen students responded that there were no challenges involved. Three students thought that it was difficult to understand the questions or information using the paper-based case study. Two made comments about not being able to "*experience it as clearly*" and that you would *"not be able to visually see the symptoms of the patient*" when using the paper version. Finally, two students made comments surrounding the lack of feedback as:

> *"Inability to correct yourself"* using the paper version: they felt there was a challenge in *"not finding out if you had answered the questions correctly".*

The responses concerning possible future uses of VR simulation at BU were diverse, and less easy to tabulate. Of the survey respondents 17 described possible future uses of the VR exercise. Three suggested including a choice about how to navigate the avatar, e.g., using the mouse as well as the arrow keys on the keyboard. One LT also recommended this during the FGD.

Fifteen FGD participants suggested developing more scenarios, e.g. community settings.

Three FGD respondents alluded to the notion of unit planning, as they discussed at what point the simulation would be used within a teaching and learning sequence, recommending that it could be positioned between a first lecture and follow up seminar. Participants (n=2) also recommended adding a chatbot feature and they identified the potential to include more valuable feedback via the addition of videos, teaching and instruction. Finally, one FGD participant discussed how the simulation could be useful for other healthcare professionals, such as Healthcare Assistants:

> *"On the HCA training I can see it being utilised. It would help to show HCAs when to use machines for obs, how to use them and when to inform the nurse in charge. I can certainly see enhancing HCA training, basic core skills on a ward."*

6.4 Themes Drawn from the CR Review and the Conceptual Model

Qualitative comments will now be reported using the four themes (immersion, knowledge, confidence, and engagement), which were latent variables in the conceptual framework developed using the CR review of the literature. The theme "experience" has not been included here (because it did not naturally occur within the FGDs), but it will resurface as part of the discussion.

Comments relating to engagement fell into two subcategories, namely "enjoyment" and "ease of use". These two subcategories will now be discussed in turn. A recurrent theme in the interviews and open-ended survey responses, from all groups of participants (e.g., LTs, Adult & Mental Health Nursing students and the academic involved) was a sense amongst participants that the VR software was engaging to use.

Of the 25 participants who discussed enjoyment of the VR exercise, seven students reported that it had been “fun” to use. Moreover, several students (n=6) felt that it was more interesting than their usual ways of learning, e.g., through PowerPoint slides. Several participants (n=5) mentioned the gamification element of the software. Four felt that it was a very interactive way to learn. Finally, some students (n=2) reported that it was a “*good starting point*”. These viewpoints all point to the fact that the VR intervention was deemed to be more interactive and fun to use than normative methods. This reflects the literature, and gamification to engage learners. For example, in a systematic review of RCTs (n=30) conducted by Gentry (2020), the overall finding was that nurse education methods that used gamification, produced satisfaction at least comparable to normative teaching methods. It was expected that the VR simulation would be enjoyable to use as it was a novel idea compared to normative teaching methods at BU.

Focusing on the second element of engagement, many participants (n=26) discussed how they felt that the laptop accessed (non-immersive) VR was easy and intuitive to use, quick to download (n=3), and straightforward to navigate the avatar through. A common view (n=3) amongst LTs was that students “*picked it up pretty easily*”. Transferability of skills students learnt through using this specific VR diabetes simulation and possibilities for distance learning were highlighted by others (n=8). A few students (n=2) mentioned that they had not required any support from LTs to complete the exercise.

A striking qualitative finding in relation to the engagement sub-theme of “ease of use” was the fact that participants (n=4) commented that they found the VR simulation “*quick and intuitive to figure out*”, even if they had not played computer games in the past. In this way the theme of experience was indirectly alluded to. This concept will be revisited within the discussion of the findings, because it supports the notion that for VR technology to be effective there is no need for users to have prior computing or specific VR experience, and this is further evidenced in the literature, e.g., Falconer (2017).

Overall, out of the comments that related to “ease of use” (under the engagement theme), the most surprising positive views related to how helpful the inclusion of text boxes had been. Text boxes were a feature of the VR diabetes simulation because more advanced Chatbot capabilities would have been more costly. Text boxes were used in the simulation in three main ways: to provide a nursing handover in order to inform the student nurse avatar of the patient’s care so far, to pose clinical dilemmas to the student nurse avatar in the form of MCQs, and finally to provide feedback to the student nurse about how effective their choice had been in terms of diagnosing or treating the deteriorating patient in the scenario. For example, text boxes were used for the nurse in charge to give the handover to the student nurse avatar. There was an unexpected advantage of using text boxes in the VR simulation, which some students (n=5) felt aided their learning. A further advantage of the text boxes (n=2) was the chance for repetition that they provided:

> *“I found having the text box there means you can read it more than once, so you do not have to recall what it said.”*

When investigated further as part of the FGDs, students (n=4) tended to believe that a combination of audio and text boxes would be beneficial in future iterations of the VR software, particularly in terms of inclusion of students of varying learning needs, including dyslexic tendencies. The effectiveness of the visual and audio techniques in HE, particularly for students with ALNs, has been evidenced by research (Hanks and Eckstein, 2019). Ali et al. (2020) found that nursing students with dyslexia actually preferred learning via real-life problem-solving scenarios (such as the one reported in this thesis) compared to passive learning, which is often the default mode, in traditional lectures.

Not only was repetition viewed as being an advantage of the text boxes, perhaps more importantly, they were found (n=5) to aid recall of complex drug names. Pharmacology is often perceived as being a challenging topic for undergraduate nursing students (Preston et al., 2019). Indeed, the complex nature of pharmacotherapy, high rates of medication errors, and the expanding scope of the practice of registered nurses, particularly in relation to pharmaceuticals, are all viewed as challenges (Keijsers et al., 2014). Gill et al. (2019) specifically recommended the use of gaming, simulation and online teaching formats for pharmaceutics, after they conducted a systematic review of the literature (n=20) which examined the teaching of pharmacology to registered and student nurses. The deteriorating patient case study evaluated via research reported in this thesis contained several tricky drug names, including Gliclazide and Sulfonylurea. Such spellings could prove difficult, particularly for those who have dyslexia, a learning need which can often manifest in short-term memory challenges, and difficulties not only in acquiring vocabulary, but also in pronouncing new words. A combination of both audio and visual prompts would be beneficial for such students. A few students raised the notion of how they felt VR technology (such as that evaluated through this thesis) could help them to overcome some of their additional learning needs (ALNs). The following comment illustrates this point:

> *"I do not have a science background. This exercise made it much easier for me to visualise and understand."*

Related to the discussion of the use of text boxes and potential audio features, were comments made (n=12) relating to how the "visual" element could aid learning. The fact that it is "interactive" and "experiential" was also highlighted by eleven participants:

"It had a bit of a flow to it as well, so you would speak to the nurse in charge then you would go to the patient and take his blood glucose; that was good. Also, for a second-year nursing student who has never been in a hospital before, can you believe it? It gives a lot of context to what I would actually be physically seeing, which is actually very useful."

The visual and experiential components of the VR simulation were discussed in both Chapters One (the background) and Chapter Two (the systematic review of the literature. Such features are part of the reason why the VR simulation was more efficient at improving short-term learning gain when compared to a traditional lecture. The props and avatars within the VR environment brought the learning to life and not only aided learning within this diabetes scenario but also activated (and helped to translate) previous learning for students, particularly those who had not been on an actual ward up to this time.

Another aspect of the engagement theme that also relates to learning theories and learning approaches is that of personalised learning. Within normative instruction for undergraduate nurses, students are often required to work as part of a small group. Two participants discussed how they felt that using the VR simulation aided personalised learning, in which they could make mistakes and not be swayed by a group decision. They felt that they were often asked to respond to questions and activities as part of a group which meant that sometimes some learners switched off and some just gave the same answers as their friend. As an example, one respondent commented that she felt that:

"VR would be the best way to learn from a situation as you are able to think for yourself and make your own decisions without influence from others, such as skills lab."

Moving on to the theme of immersion, overall, the theme was referenced on 40 occasions during the FGDs and open-ended survey questions, and it was the second most cited theme after engagement. Immersion as an advantage was discussed by all but one of the FGD participants. Many of the views of the participants (n=7) were about “making it real” and typical opinion was captured in this comment:

> *“It combines the advantages of a simulated ward with those of a drama role play, in that you can have the sense of urgency as well as deteriorating vital signs in the patient. If the patient is getting anxious and you need to take their blood pressure you could make your … (avatar) talk to the patient to calm them down. It really does make you really focused on the situation and it makes it feel a lot more real.”*

The fact that the simulation was viewed as feeling “real” and immersive is important as there is a need to provide student nurses with scenarios of deteriorating patients, so that they can have an ongoing and repeated safe practice of identifying and treating such patients, whilst at the same time experiencing the sense of urgency and pressure that they would in a real life ward clinical situation.

This comment encapsulates other views (n=50) that the simulation was detailed and immersive:

> *“It was, visually it was very accurate. It was well timed with walking forward and knowing that you were addressing the patient. I liked it; yes, I liked the detail of the blood glucose machine, and the obs machine. And the fact that the nurse just actually stands still.”*

Moreover, some students (n=5) felt as though they were thinking and acting in a different way when they were in the VR simulation. They referred to the realistic nature of the virtual props, e.g., the blood glucose monitor and the blood pressure machine, and discussed how their thinking was like when they are on clinical placements. These types of comments that were related to the theme of immersion indicate that VR simulation could be one solution towards closing the theory-practice gap described in Chapter One of this thesis. Moreover, it indicates that use of VR simulation can act as a bridge between normative instruction (in this case a HE lecture) and clinical placement. Furthermore, the academic (who agreed to participate in an interview for the purposes of this research) linked the closing of such a gap to the theme of confidence as well, in her comment:

> *“If their theory-practice gap has been closed they will begin to feel more confident.”*

As has been presented thus far, immersion and engagement were the two main themes that emerged from the CR data analysis, with both the confidence and knowledge themes less frequently alluded to by participants. Overall, 22 responses related to the theme of confidence, of whom nine believed that the instant feedback and reinforcement of learning would improve confidence:

> *“I think it would help my learning because it would make me think about the options that might be available. And make me work through why I would pick an answer. What would be good and bad about each step. That would improve my confidence, I feel. Yes, it would.”*

This comment implies evidence of students moving from lower order thinking (Bloom et al., 1956), to higher levels, including understanding, analysing and evaluating, for example. This is important because students not only need to be able to recall the signs and symptoms of a condition (in this case hypoglycaemia) but moreover, to quickly analyse and evaluate information in order to make quick, accurate and safe clinical decisions (Campbell et al., 2019). Participant comments (n=5) indicate that the VR simulation tested in this thesis was able to ease learners out of their comfort zone and into the ZPD (Carbo and Huang, 2019); it ensured that learners were asked to understand their rationale for clinical decisions; and it made feedback available, in difficult situations, as necessary. It appears that encouraging students to unpick the thinking behind the clinical decisions they make is somehow linked to their subsequent confidence (and perhaps competence).

Linked to the theme of confidence was the perception that using a VR simulation to practise clinical diagnostic and other nursing skills, was the sense of feeling that the VR ward was a “safe” environment. For example, several students (n=7) felt that it was a less pressurised, safer environment, in which to learn than being on clinical placement. Furthermore, their confidence improved because they knew that they would not get into trouble if they got it wrong. It is worth noting that during this implementation of the VR simulation the analytics feature of the software was not utilised. In future implementations, academics might want to gather data about student performance (when using the VR simulation). Such use of analytics could influence confidence levels, presumably in a negative manner. Without the use of analytics, or indeed a clinical supervisor watching over them, students felt that this “safer” way to practise would aid confidence:

> *“It is the reinforcement of that learning, and the real time feedback, safely, if they are getting it right or wrong. That is the beauty of the computer version, because you can make lots of mistakes and you’re not affecting anybody are you?”*

It is clear to understand how such perceived “learning from their own mistakes”, combined with real time feedback, would boost student confidence in their learning. It would be difficult to mimic such advantages through normative and large-scale instruction, particularly if that instruction were not supported by technology.

Another advantage cited (n=7) in relation to the theme of confidence, was the repetitive element of the VR software. Students felt that the ability to complete the VR scenario multiple times, either at university or at home (particularly prior to clinical placement or before an examination), would directly influence their confidence levels. They also felt that it would improve their confidence when performing clinical decision-making within similar deteriorating patient case studies either at university or out on clinical placement. This might indicate some degree of generalisation can be inferred from these findings.

Moving on to the final theme (identified via the CR review of the literature) of knowledge, there were 19 relevant comments made in total. Fifteen participants expressed that the exercise would/had improve/d student knowledge. The academic involved expected the use of VR technology would help to translate or “transfer” the knowledge or theory to their practice. Once again, typical student and LT comments related to how “seeing” the avatars and props played a major part in improving their understanding and knowledge of the concepts involved, through targeting “different ways of learning.” Furthermore, students (n=6) felt that they had “taken more in” during the VR learning session than they had in any other previous lecture. Indeed, four students referred to linking theory to practice in their responses. One student’s comment was very insightful, as not only did they explicitly state that the VR simulation linked theory to practice, but they went on to add that the undergraduate nursing students do not normally: “get a chance to do that (link theory to practice) until we are in placement.” This is significant because it demonstrates two notions.

Firstly, it reveals that this student (and perhaps others), do not consider that normative instruction (large-scale lectures- usually reliant upon PowerPoint slides) aids the connections between theory and practice. Secondly, this student believes that by using VR technologies within the university setting, the perceived time when the theory-practice gap is traditionally closed (usually during a clinical placement block of learning) is expedited. This might suggest that if VR technologies were used to regularly complement normative teaching methods, clinical skills sessions and clinical placements, the theory-practice gap might be bridged more quickly. This might also have an impact on nurses’ competencies once they have qualified. Indeed, two students recommended that VR technology should be used in a complementary way and not in order to replace any normative modes of nurse education.

Reinforcement of learning emerged as a sub-theme within the overarching knowledge theme. Seven students discussed reinforcement of learning, of which this comment is representative:

> *"It is nice when it is used in a lecture theatre. You can see it in action what it should look like and then it is nice to know that you could go away and practise that information to reinforce what you have learnt when you get home… It gave you a choice of your actions, you can learn, then next time you can do it differently… I think it would help my learning because it would make me think about the options that might be available. And make me work through why I would pick a answer. What would be good and bad about each step."*

Two students specifically highlighted the experiential or kinaesthetic element as being the point of action for reinforcement of the learning, in their comments. For example: "I think the learning has gone in a lot. Actually, doing that, it does make you think." Moreover, the VR simulation was felt (n=4) to be particularly helpful for complex learning, as illustrated by the following remark:

> *"I thought that was very informative, the fact I can still remember is very good. And likewise, the whole thing was reinforcing what I had understood previously. ….Learning about the glucose tabs and insulin given with food, but food is not your first port of call, and even your medication given, I am familiar with Metformin but Gliclazide was new to me. The Gliclazide is something I will probably recall now having used the interaction on computer version."*

6.5 Main Findings of the Qualitative Data

The results summarised in Table 6-4 will be explored and assimilated with the quantitative findings within the Discussion Chapter (Chapter Seven).

Table 6-4 Summary of Qualitative Findings

- engagement was the most frequently cited advantage of the VR simulation, of which enjoyment and ease of use were identified as two subcategories
- VR simulation of this kind can provide opportunities for personalised learning
- the textbox feature was reported to aid repetition and recall (including recall of complex medicine names), viewed as particularly useful for students with ALNs
- VR simulations of this type were useful in a variety of settings and at different points during the academic year
- immersion was thought to not only aid learning within this diabetes scenario but also activated (and helped to translate) previous learning and aid future learning
- frequent use of VR learning, in conjunction with clinical skills teaching and lectures, was recommended in order to close the theory-practice gap (and at an earlier point in students learning)
- it was viewed that the instant feedback provided using VR technology would improve students' confidence
- VR simulation provided students with a less pressurised, safer environment, in which to learn, compared to clinical placement
- participants considered that the exercise improved student knowledge
- if VR deteriorating patient simulations are used to regularly complement normative teaching methods, the theory-practice gap might be bridged more quickly
- Despite the numerous positives cited, challenges were also raised, of which the two main challenges were software instability and avatar navigation issues.

6.6 Chapter Summary

This chapter reported and assimilated the qualitative findings from both the surveys and FGDs. The student participants reported that the VR exercise aids understanding of the complex concepts associated with hypoglycaemia, provides immediate feedback to students about their clinical decisions, can be completed multiple times, e.g., for revision/distance learning, aids visual learners, complements ward and simulated ward experiences, and finally that it provides more opportunities for safe practice. LTs and the mental health academic also cited benefits of the computer version of the exercise; however, challenges were also reported. Taken together, the results in this chapter suggest that the VR simulation was perceived as being an enjoyable and effective way of learning. Chapter Seven will now examine these findings in more depth along with the quantitative results.

Chapter 7 Findings and Discussion

7.1 Synopsis and summary of learning so far

As highlighted in the literature review, VR simulations are increasingly being used in HE to innovate teaching (Bayram and Caliskan, 2019). Several reports have shown that such simulations enhance learning (Allcoat and von Mühlenen, 2018), and provide authentic learning opportunities (Wu et al., 2013). A strong relationship between VR learning and engagement has been reported in the literature (Parong and Mayer, 2018). Whilst negative outcomes have been less frequently published (40% of studies), technical difficulties have been detailed by some previous researchers (Fernandez, 2017).

Much of the research up to now has been descriptive in nature and the methodological quality of the few quantitative studies was inconsistent. (Ogbuanya and Onele, 2018). Where RCTs have been conducted, sample sizes have not always been optimal. Previous studies evaluating VR simulators in HE observed inconsistent results on whether such training exercises improved learning outcomes (Parong and Mayer, 2018). This study is one of the first of its kind (as far as is known), that compares a VR clinical case study with a paper-based case study. The use of PLS-SEM to analyse an RCT via a CR approach is also thought to be unique to this research and it is argued that they provide holistic and insightful findings that will be useful for those wishing to optimise the use of VR to aid student learning in HE settings.

This chapter synergistically presents findings in relation to the main aim and objectives of the thesis. A comparison is made between the findings of the present research and those of previous studies. Particular attention is drawn to unexpected outcomes and possible reasons for these unusual findings.

The chapter closes with a reflection upon the CR synthesis of the research. The main findings will now be discussed chapter by chapter before moving onto a presentation of the findings in relation to the aim and objectives of this thesis. Learning taken from each chapter will now be presented as a way of recapping what has been learnt up to this point in the thesis.

The first chapter presented the finding that whilst VR has consistently been proven to be a more engaging and enjoyable method of learning for, HE students, previous research has been inconclusive in determining whether VR (either non-immersive or immersive) produces improved student learning outcomes. This first chapter also stressed the importance for VR use in nurse education to be evaluated in terms of what works and what does not work so effectively. Chapter one also concluded that if a learner needs to bring a learning environment into their traditional lecture theatre then VR is usually more suitable, because it substitutes the physical world with a computer created one. Hence VR was chosen as the technology to enhance learning for the research reported in this thesis.

Chapter Two, through systematically reviewing the literature, found that whilst evaluative measurements have become more objective, stronger evaluation is necessary to confirm the efficacy of VW/VR/AR use in HE, particularly in relation to the use of pre-testing or control group comparisons in VW research. Longitudinal, large-scale, experimental studies were also recommended to evaluate these technologies more holistically. Hence, a larger scale (n=171) experimental study was designed and implemented to evaluate the use of VR with undergraduate nursing students.

In Chapter Three, the CR literature review approach to illuminating the efficiency of a VR intervention, involved revealing outcomes in terms of a permutation of contextual mechanisms, programme mechanisms, and agency. Four key themes emerged from the review, namely: confidence, experience, engagement and immersion. After CR analysis, two CPMAOs (contextual mechanisms (CM) + programme mechanisms (PM) + agency (A) = outcomes (O) were produced. This chapter set out the themes that needed to be tested using the novel evaluative method.

For CPMAO#1 the CMs investigated were the amount of diabetic nursing experience and relevant computing experience required to complete the VR exercise and understand complex diabetic concepts. PMs included the affordances of VR including opportunities to practise via experiential learning, and affordances of a computing exercise, e.g. the immediate feedback given and the chance to repeat the exercise. Agency (A) was a potential boost in confidence related to the instantaneous feedback received by students. The outcome (O) was improved knowledge of hypoglycaemia.

For CPMAO#2 the CM identified was participant engagement level with the learning content. PMs included the affordances of VR, for example, visualisation of abstract concepts. Agency (A) was deemed to be participant immersion with the desktop VR. Finally, the outcome (O) this time was enjoyment of the learning activity. These two CPMAO configurations shaped the conceptual framework of interrelated factors that, it was hypothesised, contribute to knowledge when learning with VR. The methodology of the qualitative and quantitative aspects of study that was designed to test this theory was then presented in Chapter Four.

Chapter Four was a significant chapter because it provided justification and reasoning behind, and design details of, pairing CR with PLS-SEM to analyse the data from an RCT. The commonalities between the philosophical underpinnings of PLS-SEM and CR were discussed along with the concepts of mechanisms and variables. This discussion will help future researchers who are considering using this novel evaluative method. Details about how the VR simulation worked were also provided in this chapter.

The results of the quantitative aspect of the research (Chapter Five) indicated that students reported that when learning with the diabetes VR simulation it was easy or moderately easy to use. They deemed it to be the second most efficient way of learning after skills lab and ahead of the normative HE methods. Overall, the results evidenced that there was a difference between the Control and Experimental Groups across the PLS-SEM model. Across the groups' participants', prior experience was consistent between groups (as we would expect). However, a striking result is that having completed the case study the Experimental Group had higher mean values for engagement, confidence and knowledge. The point of action for the pathway model appears to be the engagement to immersion pathway, which was found to be statistically significant (as measured by the permutation test). Together these results provide important insights into how use of VR learning compares to traditional HE is teaching methods.

Chapter Six explored the qualitative and CR angle of the data set. Through Chapter Six, the qualitative results demonstrated that students found that the VR exercise aids understanding of the complex concepts associated with hypoglycaemia, provides immediate feedback about their clinical decisions, can be completed multiple times, e.g., for revision/distance learning, aids visual learners, complements ward and simulated ward experiences, and finally that it provides more opportunities for safe practice. LTs and the mental health academic also cited benefits of the computer version of the exercise; however, challenges were also reported.

Within Chapter Six, all participant groups highlighted the "immersive" and visualisation aspects of the VR simulation which were viewed as strengths.

The most cited advantage of the simulation was “ease of use”, and these answers included comments about being able to repeat the exercise. Stability of the software (e.g., concerns that software would crash) was raised by both LTs and students, and was the main concern about the VR simulation. The ability or ease of use to move the avatar around the virtual ward room was the second most prevalent concern amongst LTs and students. LTs were more likely to cite possible challenges than students. However, overall results in this chapter suggest that the VR simulation was perceived as being an enjoyable and effective way of learning.

Through CR, Chapter Six also pulled together responses to the question: “To whom, when and in what circumstances does VR simulation make a significant difference in HE teaching and learning settings?” Concerning “where” students would like to use the VR simulation, results showed that the VR simulation can be used anywhere (unlike an AR version) and therefore could be used at home, whilst in the skills lab or in the clinical placement. Students expressed an interest in both using the VR learning tool as part of a university teaching session and at home. This advantageous affordance of VR was also recognised by the LTs. Regarding “when” the VR intervention should/could be used, students expressed a desire to use it at various points in the academic year including prior to the diabetes teaching session, during the teaching session and prior to an examination or commencing a clinical placement.

In respect of “in what circumstances” does VR simulation make a difference to teaching and learning in an HE setting, a deteriorating patient case study involving complex physiological and pharmacological concepts was deemed to be a suitable subject matter for 2nd year nursing students. However, other such conditions where clear clinical guidelines are available might also be suitable, e.g. the treatment of an acute exacerbation of asthma.

Finally, if we now turn to the notion of “whom” is the use of VR simulation useful for, the synergistic results of this research have indicated that second year nursing students, regardless of their age, prior nursing and prior computing experience, would benefit from the intervention within an HE setting. The results indicated those who had higher immersive tendencies were more likely to engage well and learn successfully from the VR tool.

Overall, the VR simulation was found to be an inclusive learning method and particularly for those with ALNs. However, as with any VR learning exercise, it would not be suitable for those with extreme motion sickness, migraines, and visual disturbances. Having synthesised both qualitative and quantitative findings of the present study, the recommendation for academics would be to design VR environments that are enjoyable and motivating by creating a high level of usability (Makransky et al., 2019) for the VR features, which engage students and provide them with a sense of immersion.

The main aim of the thesis was: to assess the effectiveness of CR paired with PLS-SEM as a method to evaluate the impact of VR simulations on undergraduate nurse education. The aim was achieved through two objectives, which were: to determine the effect of pairing CR with PLS-SEM as an evaluative method and to determine how using this novel evaluative method can inform our understanding of the impact and future use of VR simulations for undergraduate nurse education. This chapter will now respond to this main aim and objectives via three main areas, the first of which is ‘what has been learned through creating the novel method used in this thesis?’. The second area for reflection concerns what was found when the novel method was applied to VR diabetes undergraduate nurse education and what it enabled us to find out about students’ learning? The final area concerns the question: What did the application to VR diabetes learning reveal about the method?

7.2 What has been learned from creating the novel evaluative method?

This section relates to the creation of the novel evaluative method, i.e. the two main areas of learning related to the compatibility of CR with PLS-SEM philosophy, and the use of CR reviews to construct a PLS-SEM pathway (or conceptual) model. These two areas will now be discussed.

7.2.1 There are commonalities between the philosophical underpinnings of CR and those of PLS-SEM

Firstly, through the creation of the CR/PLS-SEM evaluative method, a contribution to knowledge was made, in what is believed to be one of the first attempts to align philosophical underpinnings relating to mechanisms (CR) with variables (PLS-SEM). This thesis argues that they are philosophically similar enough to be considered as comparable. Moreover, the configural approach of PLS-SEM is echoed in what Danermark (2002) believes CR sets out to examine:

> *"what combinations of objects, what forces and mechanisms exist, and how they together contribute to the building up of the concrete phenomena at hand."*

Consideration of the philosophical foundations of CR and PLS-SEM, along with an attempt to consolidate terminology and workings that are interchangeable and comparable between the two methods, will assist researchers when they consider if this novel paired approach will be suitable for their research.

7.2.2 CR literature reviews can be used to inform and construct a PLS-SEM pathway model

Secondly, through creating the novel evaluative method, it was found that CR was a natural fit for reviewing the literature to provide evidence for each latent variable for the pathway model. Chapter Three highlighted that VR learning is highly complicated in nature as it incorporates many overlapping and related mechanisms, of which researchers have not agreed how, or in what ways, they work to influence student learning. This evidenced sound justification to pursue a PLS-SEM approach to the analysis of the effects. Furthermore, despite the prospect that different people will benefit from VR, in different amounts, there was found to be sparse research about the types of individuals who would most benefit and few specific interpretations of how VR has helped to improve the experience of learning and in what particular circumstances.

7.3 What has been learnt from the application of the novel method to VR diabetes simulation?

The second section of the findings concerns what was found when the novel method was applied to VR diabetes undergraduate nurse education and what it enabled us to find out about students' learning. Application of the novel method enabled four key areas: the identification of group differences, identification of the key point of action of VR, richer insights into the relationships between CR themes and finally, the notion of "who is VR useful for, when, where and why?"

7.3.1 Identification of group differences

Before discussing how the novel evaluative method enabled identification of and explanation for group differences, it is first relevant to discuss how the pairing method used enabled analysis of how groups were the same or similar (invariance). This is important in terms of invariance testing. In relation to group differences, it was essential to evidence that the Control and Experimental Groups were similar in terms of students' prior computing, nursing and diabetes experience. A finding that is important to keep in mind is that, although the RCT was not deemed to be a matched control, prior experience was found to be invariant between groups. The results of the scalar invariance testing (MICOM Step 3) showed that experience had equal mean values. This means that across the Control and Experimental Groups' students' prior experience was consistent. This result was predicted and hoped for because it means that the Experimental and Control groups were equally matched in terms of their previous diabetes nursing experience, experience of computing (and social media usage) and their mean ages. The fact that prior computing experience did not impact on learning outcomes when using the software is evidence of the VR simulation being an "inclusive" learning approach (Collins et al., 2017).

In relation to prior computing experience, the experience to knowledge path coefficient was found to be insignificant. This indicates that a student's previous computing usage did not impact on their learning gain when using the VR simulation; again, this signifies the "inclusive" nature of the software used in this study. The finding that over 80% of students frequently use a smartphone indicates that future iterations of the VR training program could be accessed via such devices (provided it was available on the cheaper models that students are likely to own). The results of the JISC 2018 student digital experience insights survey (cited (Matt, 2018) revealed that about eight in ten students used a smartphone to support their learning. The JISC survey also evidenced that many students bring their own devices to their university but cannot use these to access subject-specialist software. This indicates system rigidity, and lack of interoperability. Indeed, in the present study, a lack of compatible mobile phone ownership amongst students dictated the VR-desktop approach used. Another interesting finding concerning student digital usage was that although VR and gaming activities amongst 2nd Year students was very low, one quarter of students reported very frequent use of AR. This might be due to prolific use of simple and available apps such as Snapchat amongst young adults (Alhabash et al., 2019).

In terms of prior diabetic nursing experience, whilst just over half had cared for a diabetic patient, few had either been on placement with a Diabetic Nurse Specialist or on a dedicated diabetes ward or unit. Though this would be expected as the students were only in their second year of training; however, moving forward with nurse training, such specialist experience would be advised when considering the proliferation of diabetes in the UK.

Whilst prior experience was not found to be significant in contributing to knowledge gain using VR in this study, the findings showed that as approximately one third of students did not have a GCSE level science qualification this might remain as a point of consideration. Over two thirds did not have an A Level in science. The ramifications of this would suggest that students could struggle with the more scientific elements involved in the nursing curriculum, for example, pharmacology, immunology, genomics, and so forth, and complex physiology (for example, diabetes). Whilst most Nursing Degree programmes request two to three A Levels and recommend science subjects, they are not compulsory options. All Degree programmes request Grade C or above (equivalent Level 5) in a Science GCSE. As the New Nursing and Midwifery Council (NMC) guidelines imply that all nurses will now enter the profession “prescription ready”, it is important now more than ever that student nurses can understand pharmacology and related physiological processes. Under the New 2018 NMC guidelines, as part of the pre-registration proficiency standards, student nurses should receive: early access to prescribing programmes after registration, gain more knowledge regarding **prescribing practice, pharmacokinetics, pharmacology, and whole systems assessmen**t (NMC Standards for Nurses 2018 cited by Leigh and Roberts 2018). With this rising of the bar concerning the levels of expertise and skills required for newly qualified nurses, "inclusive" learning tools, such as VR, it is argued here, are needed to ensure that students are able to meet these high expectations regardless of prior experience and education.

Having discussed how the novel evaluative method enabled analysis of invariance, discussion will now turn to how the method enabled identification of variance (or difference) between groups. The evaluative method enabled analysis across groups. The key findings of the Post-test Survey Findings across groups (both Control and Experimental) was that overall, students: improved their confidence and short-term knowledge scores post intervention (either paper-based or VR-based), believe that the VR learning would be more enjoyable than the paper version, perceived that VR learning would be particularly useful for the comprehension and application of diabetic nursing concepts and finally, students felt that VR simulation would be the second most effective way to learn about diabetes after skills labs learning.

Across the data set (post-test survey) all pathways were found to be significant. This means that the experience (including prior computing and diabetic nursing experiences) to knowledge pathway was partially mediated by confidence. This means for all the participants involved in the study, the more previous experience they had, the more confidence they had when using either diabetic intervention (either paper or VR) and the higher their knowledge scores post intervention. The upper section of the diagram also shows in red (see Figure 7-1) that across groups, confidence and knowledge scores were higher in the Experimental Group, post intervention.

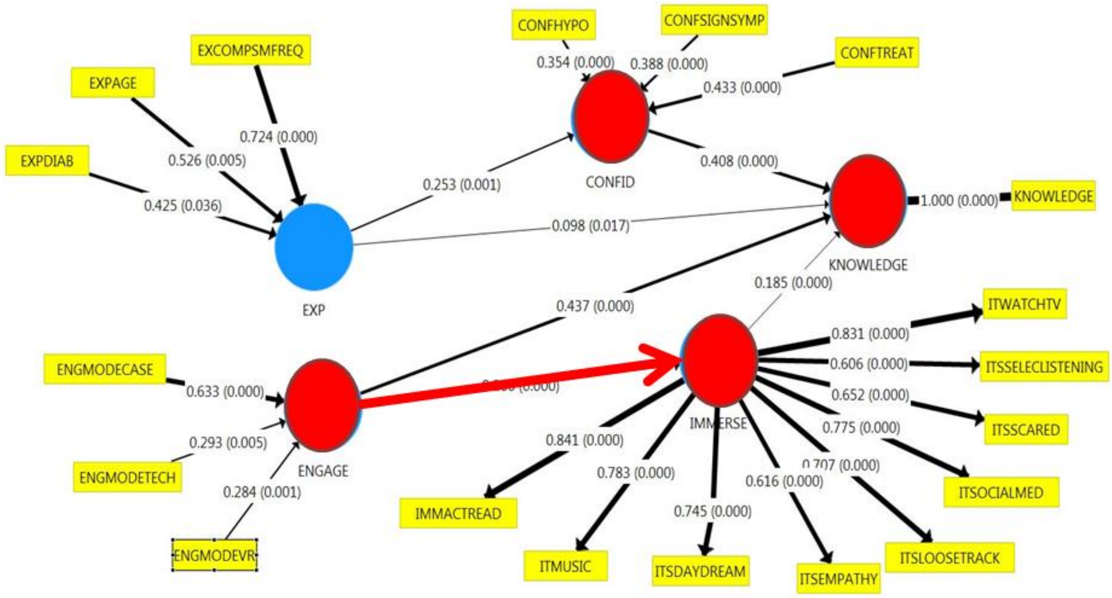


Figure 7-1 Group differences shown on the Conceptual Model

7.3.2 Identification of the key point of action of the VR simulation

The novel method enabled the identification of the key point of action; this is because the method permitted not only analysis of the differences between groups (in relation to the variables) but also across the entire pathway model. This is important because it provides details about the relationships between variables and the strength of those relationships and knowing the precise details can guide researchers and educators who are developing and using VR in HE.

Figure 7.1 shows that, in the lower half of the diagram, across the data set the engagement to knowledge pathway was partially mediated by immersion, and again these were significant pathways across the whole data set. Across groups however, the only significant pathway in the conceptual model was the engagement to immersion pathway (shown as a red arrow). This might indicate that the improved knowledge scores for the Experimental Group have resulted from this “key pathway”, which could be deemed to be the “action point” of the PLS-SEM model. In other words, for the Experimental Group, whilst learning about diabetes using the VR simulation, students became more engaged in their learning, and therefore more immersed in their learning, than those learners who did not use VR (the Control Group); and this resulted in the Experimental Group attaining higher knowledge scores (as measured by a set of ten MCQs) than the Control Group.

The latent variable prior experience (shown in blue) was not found to statistically significantly differ between groups. This means that when using the VR simulation, a student’s prior experience, including if they had used VR before, and so forth, did not impact on their MCQ scores. This once again indicates that the VR simulation is an inclusive learning tool, regardless of students’ age, computing experience or diabetic nursing experience. It means that it is a suitable learning method for all students (provided they are not susceptible to nausea or migraines and so forth). This finding is supported by earlier work of Childs (2010) and Falconer (2017) who concluded that prior experience does not affect the ability of a student to learn and have positive experiences when using VR. This is useful for educators to know when they are developing teaching and learning materials for their students. Having discussed experience, the remaining themes will now be discussed.

7.3.3 Richer insights into CR themes and the relationships between themes

Finally, in relation to what was revealed when the novel method was applied to VR, it was found that when CR was paired with PLS-SEM, understanding about the themes (mechanisms or variables) and the relationships between them were enriched. The themes that interact when VR simulation is used for diabetes education will now be discussed, starting with the theme of engagement.

7.3.3.1 Engagement and its relationship with immersion

In relation to the first theme, engagement, the novel pairing of CR with PLS-SEM to analyse an RCT facilitated detailed insight into the theme. Several aspects were evidenced via the evaluative method including:

- validation of engagement as a variable on the pathway model
- positioning and validation (through FGDs and statistics) of the engagement theme on the conceptual pathway model
- evidence of higher mean values for engagement in the Experimental Group
- identification of a significant pathway between engagement and immersion (between Control and Experimental Groups)
- indication that immersion not only aids learning within this diabetes scenario but also activates (and helps to translate) previous learning and aid future learning.

The evaluative method employed evidenced that the coefficient pathway for engagement to immerse was found to be significant. The exact explanation behind this has not been determined; however, this finding implies that the key point of action for the conceptual pathway lies in the engagement to immersion relationship. The data point to evidence supporting the notion that the students who used the VR simulation were more engaged in their learning which led to being more immersed and hence their knowledge scores were boosted.

Beyond this thesis immersive VR data will be collected using the same training exercise this time accessed via Google Cardboard headsets. It will be interesting to see if a pattern will arise between student immersive tendencies and the engagement to knowledge pathway. The fact that the engagement to immersion path coefficient was found to be the only significant path in the conceptual model suggests that this is the key point of action of the model. In other words, when considering the superior short-term knowledge gain in the Experimental Group, both student engagement with the VR software and immersion in the learning activity were the key contributing factors. Moreover, this might indicate that for the Experimental Group the positive influence engagement has on immersion is stronger than in the Control Group. In other words, the VR simulation was more engaging than the paper version and, partially mediated via immersion, improved participant knowledge scores. These results corroborate the findings of a great deal of the previous work, for example, for Martin and Carr (2009) the notion of both immersion and engagement are tightly entwined.

Beyond the statistics, the method allowed for drilling deeper into the engagement theme. Both LTs and students recognised that this type of learning was more interactive and could hold learners' attention better than traditional methods. This is especially important for subject matter such as complex physiology and pharmacology and for students who perhaps struggle with scientific concepts and/or who have ALNs. Gamification was mentioned by some of the students as it is viewed as an engaging, possibly competitive, and fun aspect of an activity. Previous researchers have described how using gamification platforms can help support students with dyslexia who might be struggling in their education (Gooch et al., 2016). According to Cugelman (2013) gamification principles are closely linked to similar principles that have been proven to work in relation to behaviour change, such as effective motivation to exploit potential. Gamification also provides an opportunity for personalised learning.

Closely linked to the notion of engagement (though treated as a distinct aspect in this thesis) is the enjoyment of learning. The data from the present study showed that the enjoyment, as reported by students who completed the VR version of the training, was approximately 20% higher than the perceived enjoyment of using the VR that the Control group reported. This suggests that using the VR simulation was more enjoyable than students expected it to be. Once students had used the VR version, they were also more interested in learning, using VR, than the Control Group; it also sparked their interest in learning in new ways and with different technology. This linked to the responses from the FGDs as students were very keen to use different VR scenarios in the future.

Ease of use was a recurring theme within the qualitative findings. Similarly, Birt et al. (2018) concluded that there was a correlation between student enjoyment and ease of use in their study of 63 biomedical students using VR to learn about tissue engineering concepts. Regarding ease of use, remarks suggested that overall students felt the VR simulation was “intuitive” and “pretty easy” to use, suggesting that students would not be put off the first time they used the simulation and would perhaps continue to use the simulation as a distance learning tool. This is significant because in the future, it is anticipated that online and distance learning modes will dominate (Traxler, 2018). It also evidenced that the VR deteriorating patient simulation created can be used either at home or within an HE seminar or lecture context.

Another consideration about “ease of use” was regarding the comment that the simulation only took “a couple of seconds or minutes to work it out. It wasn’t difficult, just a bit initially”. Students and LTs felt that the simulation was easy for anyone to pick up. In the future, time could be planned for students to practise the VR Fieldscapes tutorial, either at home or at BU. This would help to overcome any initial difficulties in using the software. Furthermore, once several simulations become available, the advantages of using them would offset time taken to become familiar with the software.

7.3.3.2 Immersion

The evaluative method implemented enabled precise analysis not only of the key themes (or variables) but also of their related indicators. Several of the immersion indicators had the highest indicator weights, as illustrated in Figure 7-2. The following indicators had the strongest positive relationships:

- IMMACTREAD (indicator weight= 0.841)
- ITWATCHTV (indicator weight= 0.831)
- ITMUSIC (indicator weight= 0.783)
- ITSOCIALMED (indicator weight= 0.775)

This would be expected as reflective indicators tend to have higher weightings. The relationship between immersive tendencies levels, including when reading, watching television and using social media, amongst students and the rest of the VR simulation learning conceptual model, was not unexpected. This might be because the VR software incorporated reading of text boxes and viewing through a digital screen for example, hence the crossover of immersive tendencies between reading, watching television and using digital social media and using the VR simulation.

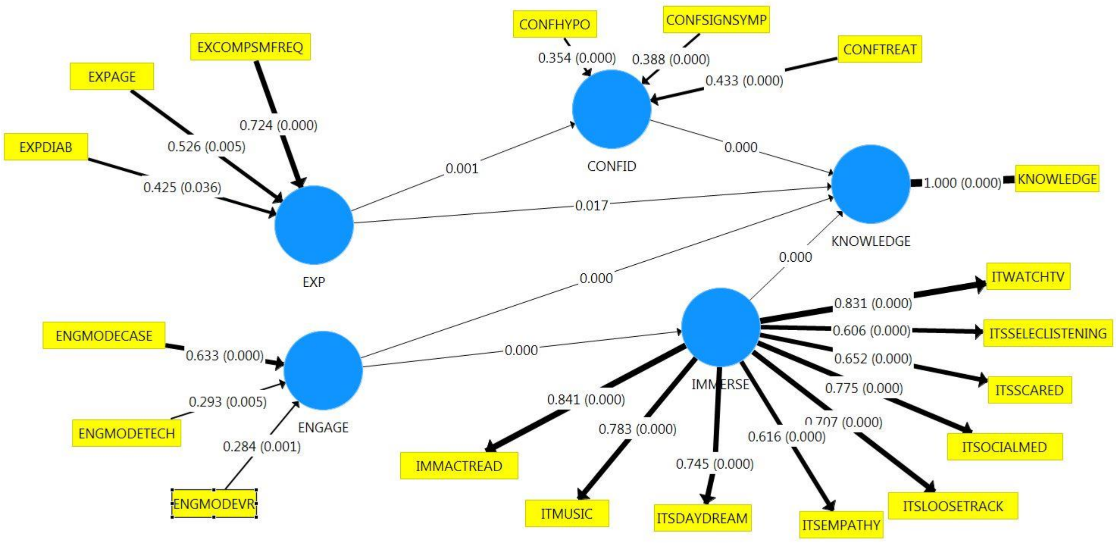


Figure 7-2 PLS-SEM Diagram showing indicator weightings (outside of the brackets).

The final point to be made concerning the quantitative findings that related to immersion was that the f^2 effect size analysis indicated that immersion only partially mediated the engagement to knowledge path. This result is difficult to explain but could point to students' immersive tendencies not being an essential requirement for their engagement with the VR learning approach, and hence suggestive of the inclusive nature of VR simulation as a teaching method. However, further in-depth mediation analysis is beyond the scope of this thesis.

In addition to these very fine statistical observations, the novel method enabled additional insight. For example, students felt that the use of VR brought the clinical case study "to life", with some remarking on its combined drama role play with a simulated, or real ward experience. Students believed that the VR simulation added a sense of "urgency" to the situation. In a high-pressured ward environment, there would be a sense of urgency and a requirement for nurses to make fast accurate clinical decisions about patient care (Horwood et al., 2018).

Some observations made by students suggested that they were "immersed" in the simulation, as they spoke of "thinking as if they were on the ward" and expecting to find various pieces of equipment where they are located at the local hospital. This reminder of ward-based nursing practice during their teaching unit at the university must go some way to bridging the theory-practice gap (Greenway et al., 2019). This was made apparent in the comment from the student who stated that she had not been on placement on a traditional ward setting yet, and that even just seeing, or "visualising" the ward and equipment (albeit in VR) was a learning opportunity for her. This suggests that other low-cost technologies such as 360-degree cameras in conjunction with the VR headset would also bring affordances, such as "visualisation of concepts" to first- and second-year nursing students. Though envisaging is one of the most obvious advantages of VR, critics might argue that this likewise can be achieved via video. Nonetheless, videos are inactive learning objects, while VR permits interaction with the environment (Allcoat and von Mühlenen, 2018). It is evident that learning through VR enables diverse approaches towards learning. Different types of information might be better presented in some formats rather than in others. For example, language may be best learnt with audio, while nursing may be more suited to visualisation. Personal user preference is also an important consideration.

7.3.3.3 Confidence

Moving on to the theme of confidence, the novel evaluative approach facilitated detailed insight into the theme. Again, several aspects were demonstrated including:

- improved (higher mean values) confidence for students who used the VR simulation
- possible reasons behind this improvement in student confidence

The evaluative method meant that not only was improved confidence demonstrated statistically but, additionally, the method revealed some of the reasons why confidence improved in the group of students who used VR simulation. Feedback was found to be a substantial and recurring theme throughout the analysis process. Students who had used the VR exercise reported that they felt it offered the advantages of the ability to learn from mistakes (55%) and repeat the exercise (51%) when paralleled with traditional learning approaches. Some participants stated that they thought the instant feedback given was easier to accept from a computer than a mentor in a real-life ward situation.

In addition to the implication that the instant feedback provided by the VR simulation contributed to the short-term learning gain in the Experimental Group, the “lack” of feedback was also cited as a challenge of the paper-based version of the diabetes case study. Considerable research has been undertaken within the field of education research that has concluded that, when used purposefully, formative feedback, and/or feedforward, aids learning (Black and Wiliam, 1998, Morley et al., 2019).

Students found that the immediate feedback provided by the VR exercise boosted their confidence when making clinical decisions within the scenario. This immediate feedback meant that students were able to change their decisions at any stage during the clinical scenario. Though this is something that might not be available or appropriate in a real-life ward setting, it has advantages in terms of aiding confidence and competency. For example, in a clinical setting it might be appropriate to change clinical decisions at certain points but not at other points during nursing care. The fact that the VR simulation had no time limit led students to appreciate that they could have many goes, make mistakes in their clinical decision-making, and yet not affect patients in a negative way. It could be surmised that this also led to students feeling confident about their learning and is an advantage compared to learning live with their mentor in a placement situation. Again, this highlights the recurring sub-theme of the “inclusive” nature of the VR software, as students with slower cognitive processing speeds (Saban-Bezalel and Mashal, 2019) would not be disadvantaged when using the program. Knowing that the instant feedback provided by these VR simulations, and the less pressurised and safer environment provided by the technology improves student confidence, will be useful information for people who are creating and implementing similar simulations with HE students.

The pairing of CR with PLS-SEM for analysis of the RCT enabled detailed and more precise analysis of the confidence theme, and which areas of confidence were affected. For example, the overall finding was that both groups' confidence about their diabetes knowledge improved after the intervention. The pairing method enabled scalar invariance testing which revealed that the Experimental Group had higher mean values for confidence than the Control Group. This means that students who used the VR software felt more confident after the intervention. Specifically, the post-test survey revealed that the Experimental Group felt more confident about explaining signs and symptoms and treatment of hypoglycaemia, and slightly more confident about how to use each item in the Hypobox, when compared to the Control Group students. The f^2 effect size analysis indicated that confidence partially mediated the experience to knowledge path. This might explain why the path coefficients were not found to be significant individually, though higher mean values were evidenced for both knowledge and confidence.

7.3.3.4 Knowledge

Finally, turning now to the theme of knowledge, the pairing of CR with PLS-SEM promoted detailed comprehension of the theme. Findings included:

- better (higher mean values) short-term knowledge gain for students who used the VR simulation
- the Experimental Group answered all ten MCQs about hypoglycaemia more accurately than the Control Group
- the Experimental Group were 11% better at knowing when to administer Dextrose, 10% more efficient at correctly identifying which oral medication can lead to hypoglycaemia and 9% better at selecting the correct initial treatment for hypoglycaemia
- improved understanding of when and how to use each item in the Hypobox for those students who learned via VR technology
- students, academics and LTs think VR improves student learning outcomes.

Therefore, it was found that VR simulation makes a significant difference to second year student nurses' knowledge when compared to learning with a paper version of the same case study. The novel evaluative approach also provided a wealth of insight into the reasons behind “for which students”, “how” and “in what circumstances” enabled this improvement in learning. Starting with “for which students”, the textbox feature was reported to aid repetition and recall (including recall of complex medicine names) and was viewed as particularly useful for students with ALNs. In relation to “in what circumstances”, VR simulations of this type were useful in a variety of settings and at different points during the academic year. Finally, in relation to “how” the VR simulation was felt to improve student knowledge; as mentioned above, the engaging and immersive (e.g., experiential and visual elements) nature of the software was felt to play a major part in creating successful learning. The evaluative approach shone a light on the personalised learning opportunities it provided and the VR simulation’s potential for not only closing the theory-practice gap; moreover, participants believed that if VR deteriorating patient simulations (of this kind) are used to regularly complement normative teaching methods, the theory-practice gap might be bridged more quickly.

The evaluative method used in this thesis enabled comprehensive understanding about improved knowledge, beyond the statistical findings. Additionally, the method provided participant views about how effective they feel using VR simulation is for diabetes nurse education. Having used the VR simulation the Experimental Group were 36% more likely to think that VR was the most effective way (compared to paper-based learning) to learn about complex concepts, including diabetes, than the Control Group. This indicates that it is difficult for students to imagine how effective and useful a VR training tool can be without having used it, despite previous experience of using a VR Dementia Walk Through app.

The evaluative method also facilitated further probing into the reasons behind “why” VR was more helpful to students than the paper case study. Specifically, LTs viewed the visualisation aspects of VR to be superior to normative learning. Similarly, students discussed the fact that they thought VR simulation would be helpful to reinforce learning via a kinaesthetic, interactive and experiential approach. Students reported that the VR version was helpful in learning the “tricky concepts” involved in long-term conditions and pharmacology.

Once again, participants (including LTs, students and the academic), discussed how they felt the VR affordances of the software would help to bridge the theory-practice gap, making it easier to “translate” the knowledge from lecture theatre to ward, via targeting different ways of learning using the VR training exercise. Students felt that they had learned more about how to treat patients with a deteriorating condition than they had in preceding lectures. The fact that the learning connection would be made at a sooner time than usual was also raised, along with the affordance that the VR case study could be repeated at students’ leisure as well as being used to refresh students’ learning prior to going out on placement or before an examination.

Another, finer detail that was highlighted was the affordance of personalised learning. All students could proceed with the patient case study at their own pace, something that has long been a challenge for traditional learning methods (Wanner and Palmer, 2018). This point was further developed by the students who discussed the fact that it enabled individualised learning. Another feature of the VR software used in the study that was perceived as aiding personalised learning was the text boxes. The fact that the textboxes that popped up to provide the questions and feedback were deemed to be useful was an advantage, in that students could re-read the information and take their time. The pop-up question and feedback boxes can, therefore, be viewed as being gamification devices in this context. This, once again, has ramifications for personalised learning and for learners to operate at their own pace and according to their own learning needs. Students expressed how much of their learning was carried out in small groups and that this sometimes led to more passive learning for some students. As students had to complete the VR training exercise on their own, they had to engage with and become active in their learning; this could aid achievement and possibly confidence.

7.3.4 Who is VR useful for, when, where and why?

The unique pairing of CR with PLS-SEM to analyse the VR simulation RCT provided comprehensive findings. Allana and Clarke (2018) recommended exploring what (intervention) works best for whom (populations), where (settings), and why (mechanisms). For this application, the “what” was a non-immersive VR diabetes simulation? The “whom”, “where” and “why” dimensions will now be discussed in turn and will also include an additional element “in what circumstances”, which was derived from CR philosophy.

Concerning “where” students would like to use the VR simulation, students showed interest (36% definitely would) in using such a VR simulation at home to repeat and practise for an exam, and so forth; this was also cited as an advantage (distance learning) by an LT in the FGDs. Furthermore, 70% (n=152) of students felt that learning about a diabetes case study in a skills lab would be the most effective approach to learning. VR learning was deemed to be the second most effective way, followed by paper-based learning. This suggests that a VR deteriorating patient simulation is useful as it is university-based but is also useful as a pre-teaching tool in preparation for students prior to a ward placement or skills lab training.

In relation to the “where” aspect, the VR simulation can be used anywhere (unlike an AR version) and therefore could be used at home, whilst in the skills lab or in the clinical placement. Students expressed an interest in both using the VR learning tool as part of a university teaching session and at home. This advantageous affordance of VR was also recognised by the LTs.

Regarding “when” the VR intervention should/could be used, students expressed a desire to use it at various points in the academic year including prior to the diabetes teaching session, during the teaching session and prior to an examination or commencing a clinical placement. Due to the nature of the VR software, there is flexibility and scope for the case study to be made available to the students at a point when the academic thinks it is appropriate, then students would be able to repeat the exercise in their own time to suit their individual learning needs.

In respect of “in what circumstances” does VR simulation make a difference to teaching and learning in an HE setting, a deteriorating patient case study involving complex physiological and pharmacological concepts was deemed to be a suitable subject matter for second year nursing students. In the scenario a clear set of clinical guidelines was available (Diabetes UK) for the identification and treatment of hypoglycaemia. Other such conditions where clear clinical guidelines are available might also be suitable, for example (though more complex and possibly more appropriate for registered nurses) the identification and treatment of: DKA (Diabetic Ketoacidosis) and HHS (hyperglycaemic hyperosmolar state), which are two main acute complications of diabetes (Gosmanov, 2018). Other potential future scenarios could include an acute episode of asthma in a paediatric patient. Clinical conditions that would benefit from this type of VR training tool in the future include sepsis and meningitis, though clearer diagnostic tools would need to be developed in the first instance.

Finally, if we now turn to the notion of “whom” is the use of VR simulation useful for, the synergistic results of this research have indicated that second year nursing students, regardless of their age, prior nursing and prior computing experience, would benefit (in terms of subsequent learning gain) from the intervention within an HE setting. The results indicated those who had higher immersive tendencies were more likely to engage well and learn successfully from the VR tool. Hence, educators would be well advised to encourage such immersive tendencies in their students. If more VR scenarios are made available to the students, and once students have used such simulations several times, they might feel more at ease and more likely to allow themselves to become immersed within the VR.

Overall, the VR simulation was found to be a relatively inclusive learning method and students felt that it would be suitable for learners with additional needs such as dyslexia, ADHD, and so forth (Miettinen, 2019). However, as with any VR learning exercise, it would not be suitable for those with extreme motion sickness, migraines and visual disturbances (though glasses wearers would be able to use the software without difficulty).

7.4 What did the application to VR diabetes learning reveal about the method?

The application of VR diabetes learning to the novel evaluative method permitted three main things. First, it enabled the novel method to be tested, which revealed the benefits and challenges of the method. Second, it facilitated comparison between previous research and the research reported in this thesis conducted via this novel evaluative approach.

7.4.1 Benefits of the novel evaluative approach

Two main benefits were identified when the novel evaluative approach was applied to VR simulation learning. The first concerns how easy and scalable the approach will be for researchers to use. The second finding relates to how the approach taken led to the construction of a comprehensive conceptual framework.

7.4.1.1 CR/PLS-SEM analysis is user friendly and scalable via SmartPLS software

Having applied the novel evaluative approach, what has been learned from its application to a VR deteriorating patient simulation? The first thing that was learnt was that the novel evaluative model (namely the pairing of CR with PLS-SEM) is straightforward and comprehensive to carry out using SmartPLS software. Insights into the limitations and benefits of the approach were also identified.

Another area of learning that relates to learning about the creation of the novel method concerns the usability of this novel pairing of CR with PLS-SEM. Indicators and latent variables drawn from the pre and post-test surveys (developed based upon the CR review) were brought together using one piece of software, namely SmartPLS 3.0. The software enabled quick specification and estimation of the path models, via dragging and dropping items from the indicator's menu onto the modelling window, to create new constructs. All data were processed and batteries of statistical tests, supplemented by bootstrapping, were performed all within the SmartPLS 3.0 software. Browsing between descriptive statistics, state-of-the-art measurements and structural model evaluation metrics was completed with ease.

The author of this thesis is a novice statistician and found performing the PLS-SEM analysis straightforward using the step by step guidance of Hair et al. (2016) and Hair et al. (2018). Expert programming skills were not needed, and the user was not required to undertake any additional mathematical algorithms or calculations. The approach enabled understanding and modelling of the abstract statistical concepts. It was evident that the visual method would be useful for most researchers, particularly those with dyslexic or dyscalculic tendencies. In short, when CR was paired with PLS-SEM, it was found that the method was a comprehensive yet user friendly way to approach a mixed method analysis of an RCT. This means that the approach will be scalable.

7.4.1.2 CR paired with PLS-SEM enables construction of comprehensive conceptual frameworks

This section progresses through the CR strand of the evaluative method and culminates in the proposal of a CR/PLS-SEM conceptual model for VR use in HE. Evaluation of an intervention must comprise deliberation of how diverse kinds of people respond to it and why (Blackwood et al., 2010). For this reason, Porter (2015) believes that human agency is a central consideration. Having conducted the CR literature review (Chapter Three), two CPMAO configurations were designed (Appendix Six).

Within this research the PLS-SEM analysis sought to shed light upon the precise nature and relationships of three variables related to this CPMAO#1 configuration. In this instance, the analysis did not confirm that experience led to confidence (when using the VR simulation) which in turn would lead to knowledge; though it did evidence that both confidence and knowledge improved post VR intervention. This supports previous findings that factors that interact when learning in VR are complex and vary from person to person (Liu and Chu, 2010). This confirms both the notion of “agency” and the powers of personal interpretations and reactions to an intervention; it also supports the findings of previous researchers (Mayer et al., 2008, Cobbett and Snelgrove-Clarke, 2016), that not all learners benefit from the affordances that it provides.

Overall, the **<u>CPMAO#1 configuration was not substantiated</u>** through the CR/PLS-SEM findings of this research. This means that whilst separate elements of the configuration were confirmed by findings, the configuration itself was not evidenced by results. **<u>This is a positive result</u>** because it means that the VR simulation designed is an “inclusive” learning tool. It can be considered inclusive because success for the user would not rely upon previous computing or diabetic nursing experience, or a user’s age or gender.

Overall, the **<u>CPMAO#2 configuration was substantiated</u>** by both the quantitative and qualitative results. This means that not only were all isolated parts of the configuration evidenced by the findings, but the entire configuration was confirmed. Hence the conceptual pathway (Figure 7-3) was produced. **<u>This is also a positive result</u>** which links the visualisation affordances of the VR simulation to enhanced student enjoyment.

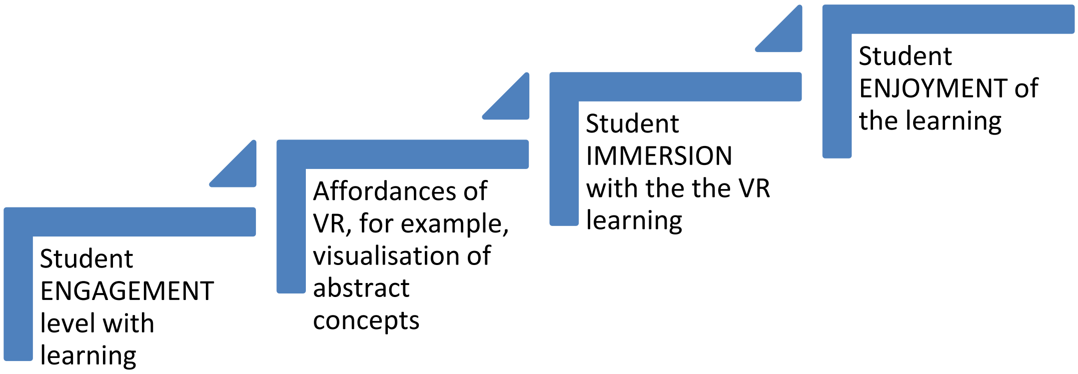


Figure 7-3 Conceptualisation of the Critical Realist Configuration Findings

If this conceptualisation is then overlaid with the conceptualisation of the PLS-SEM analysis results, we arrive at the conceptual theory shown in Figure 7-4. However, this proposed and amalgamated conceptualisation would require new PLS-SEM analysis to test if it acts in its entirety.

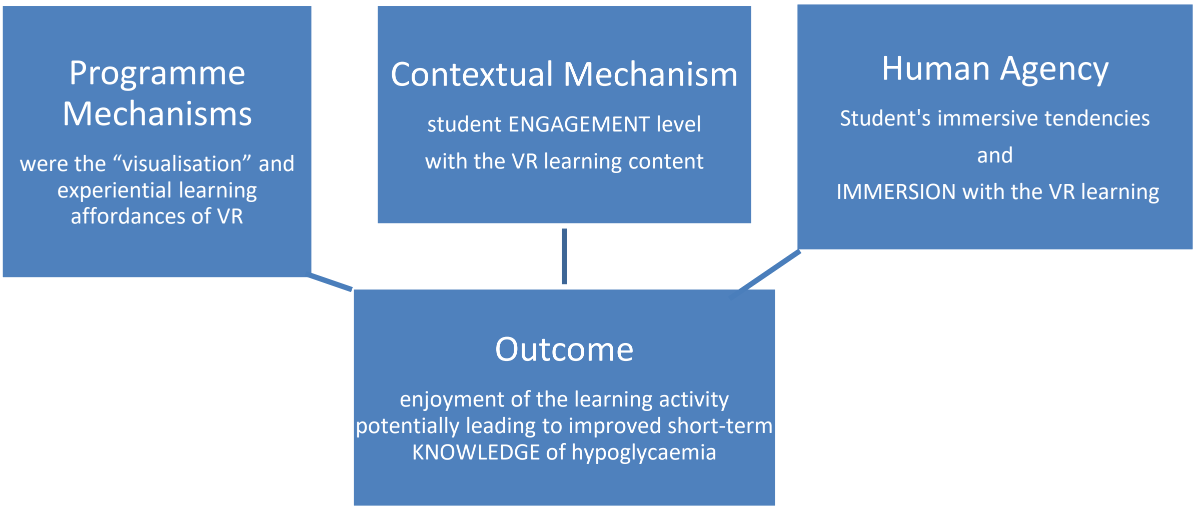


Figure 7-4 Amalgamation of both the CR configural and PLS-SEM tested models

The exact point of the action of student enjoyment remains unclear. It might be that student enjoyment acts throughout the process. The CR approach aimed to investigate what combinations of objects and mechanisms exist, and how they together contribute to the building up of the concrete phenomena taking place. It is proposed that student enjoyment requires further investigation.

The PLS-SEM approach to analysis enabled scrutiny of such combinations of the mechanisms and their exact interactional relationships. Overall, in relation to the CPMAO PLS-SEM mapping, a statistically significant indirect effect (contextual mechanism to program mechanism) was found for the second CPMAO configuration (CPMAO#2).

For the first CPMAO configuration (CPMAO#1) the contextual mechanism to outcome via program mechanism pathway was not found to be statistically significant on this occasion. Hence the relationship between experience, confidence and knowledge is more complex and nuanced than was predicted. This might suggest that confidence and knowledge do not have a linear relationship but perhaps are more cyclical or mutually reinforcing in nature. Törnberg (2018) asserted that: "*when operating in open systems, the relationship between causal mechanisms, and their effect is not fixed, but contingent*." The findings of this research uphold this view, despite inconclusive results concerning the precise relationship between some of the mechanisms involved. Törnberg's (2018) thoughts on complex causality hint at the fact that whilst a mechanism might be "necessary" it might not be "sufficient":

> *"complex causality, that is, that when operating in reality the abstracted mechanism generates the expected outcome only when operating together with another mechanism".*

One possible explanation for the finding that the first CPMAO configuration was generally found to be of no statistical significance could be that the contextual mechanisms and program mechanisms suggested were not the only mechanisms working. For example, there might have been other mechanisms working that have not been identified yet. It is indicative that some mechanisms were inadvertently overlooked, or that minor ones were included.

7.4.2 Challenges of the novel evaluative approach

The challenges of pairing CR with PLS-SEM are as follows:

- testing of this novel method is required
- full consideration of the philosophical compatibility of each strand of the method is needed
- researchers would need to follow the method in a step by step way to ensure that mistakes are not made.

It is felt that if researchers are prepared to follow several, reasonably simple, steps and procedures, then this novel evaluative method can be reproduced and scaled.

7.4.3 How do these findings compare to studies that have used other methods?

From using this novel evaluative method, when comparing back to previous research and literature, the findings tended to be similar; however, there were some differences and some areas where the novel method provided greater understanding of the themes related to using VR in, HE and its efficacy.

Via the innovative method, this research has been able to establish that VR simulations in HE can lead to improved short-term learning gain. Not only did study participants relay their experiences of positive learning when using the VR software, but a significant and measurable knowledge difference was found between groups in the current study, and this is an unusual finding. There have been some previous researchers who have reported greater conceptual and procedural knowledge learning gains following activity with the desktop VR simulation compared to the control (Chen and Tsai, 2012, Dubovi et al., 2017, Rizov and Rizova, 2015, Turkan et al., 2017, Achuthan et al., 2017, Ke and Hsu, 2015, Pellas, 2016). Birt et al. (2018) also found improved knowledge gain for a VR group of learners, but the sample was small in their study. There have been few studies that have been able to evidence measurable learning gain (Boet et al., 2013, Kirkman et al., 2014); moreover, some studies have suggested that little can be established about the retention of knowledge acquired when using VR (Mikropoulos and Natsis, 2011, Cobbett and Snelgrove-Clarke, 2016, Makransky et al., 2019). This study, however, adds to the body of evidence in relation to VR learning and retention of knowledge because the results support the notion that VR learning can improve “short-term” knowledge gain.

Through using the original method, the findings of this thesis support the findings in the literature of the significance of lifelike simulation experiences on competence and confidence (Lee et al., 2019, Doherty-Restrepo et al., 2017). For example, Su et al. (2013) demonstrated improved student confidence when learning engineering with VR, compared to traditional methods via their RCT. Similarly, Smith and Klumper (2018) found improved confidence after VR training in their pre-test/post-test study; albeit, a sample size of 15 participants limits these findings. Although this research did not find the confidence to knowledge path coefficient to be significant, the notion that confident students learn better (Zieber and Sedgewick, 2018), seems to have been implicit within the qualitative results of this study. The findings of this research suggest that efforts to develop student confidence are a substantial part of effective learning.

Within previous research and literature there has been no real commentary on the inclusive nature of VR simulations. The evaluative approach used in this thesis distilled aspects of inclusion more strongly, including instant feedback, visualisation aspects, individualised learning opportunities, repetition opportunities, the chance to make mistakes in a “safe environment”, text box information (to support students with dyslexia), and experiential learning. These elements added together enhanced student confidence for those who completed the deteriorating patient study via VR simulation, regardless of their prior experience and regardless of their learning needs.

Taken together the data evaluated via the pairing method, confirmed that engagement is associated with immersion and with improved knowledge when using VR simulation for nurse education. This finding supports those of previous researchers; for example, they compare to the findings of the distilled results of the systematic review, that VR simulation can improve student engagement (Barnett et al., 2016) and that it can lead to enriched motivation (Birt et al., 2018; Rizov and Rizova, 2015; Turkan et al., 2017). Barnett et al. (2016) found that virtual patient cases may offer limited benefit over paper cases in improving overall student self-confidence to provide medication management. However, their study was not heterogeneous and did not randomise students into experimental and control groups. Therefore, students with like interests may have been clustered in similar sections and thus may not be representative of the wider student population. The evaluation conducted in this research, however, has used randomisation and heterogeneity, and therefore the findings can be more generalisable in this instance.

Another area where the approach taken provided further insight into an aspect of VR learning relates to higher order learning. According to Bloom, cited by Su et al. (2004), application and understanding are more important than recall. In a study by Allcoat and von Mühlenen (2018), participants had reported higher learning in the VR condition for “remembering” (or recall). However, findings from this thesis evidenced that Experimental Group students felt that the VR simulation would be most helpful for them to apply

(84%) and understand (84%) the diabetic concepts they were learning about. Recall was only cited as an advantage of the VR simulation by 68% of students in the current study. Allcoat and Mühlenen’s (2018) study only looked at the lower ends of the learning hierarchy, remembering and understanding, and analyses relied upon various ANOVA tests and qualitative feedback. It is argued that the evaluative method used for this thesis was a more comprehensive approach and found that VR compares differently from normative HE teaching methods for applying, analysing, evaluating and creating.

Continuing with the theme of higher order learning, other areas of insight were obtained via using the CR/PLS-SEM approach. According to Lau and Lee (2015) accessing the various levels of Bloom’s taxonomy reinforces the learning of complex concepts. Findings from this thesis evidence that the notion of reinforcement of learning also lends the VR approach to pre-teaching of vocabulary and ideas that will be used in upcoming learning. For example, students with ALNs such as dyslexia and ADHD could practise with the VR simulation prior to receiving their Unit teaching (e.g., lecture/seminar). In this way they would be able to begin to learn complex medication names and more difficult physiological processes and feel more prepared when they learn about these in their lectures.

Previous literature and research warned that students could find learning using VR irrelevant if it was not closely aligned to learning objectives and examinations (Matthews et al., 2018). In the current study, however, this did not emerge as an issue and was not discussed either in the FGDs or the open-ended questions of the survey. This could be partly explained by the fact that the VR exercise was specifically designed to be embedded within the curriculum and was tailor-made to align with course aims. Indeed, Hack (2016) stated that technology must facilitate learning that matches the learning objectives in order to engage students. A clear connection must be made between course learning objectives and course activities when using VR (Inman et al., 2011). Several recommendations were made by Makransky et al. (2019) in terms of optimising the relevance of the VR learning activity; these were to ensure that students have a high sense of control and active learning, and that students appreciate the cognitive advantages of the VR lesson. This indicates that preparatory work should be undertaken with students not only for them to practise using VR software prior to a seminar but also in relation to their understanding about how VR can be useful to them, for example by linking what they have heard in lectures to what they have done when working on a ward as a student nurse.

Use of the original method enabled identification of an unexpected angle in relation to pedagogical underpinnings of the VR simulation, which was the finding that nursing students valued the “rarer” opportunities to work individually in HE learning sessions. Previous research tends to highlight the collaborative affordances of VR technology (Falconer, 2018; Tuzun et al., 2019). However, in the present study the chance to learn at the student’s own pace and without the group pressures to select a response that they did not necessarily agree with, was brought to the forefront. The value of this “individual learning gain” corresponds to previous research conducted by Birt et al. (2018). Hence, the use of a VR learning tool in the present research study has been viewed as a step towards personalised and student-integrated learning.

Within future iterations of the VR software, the goal will be to add in some attributes of Connectivism along with the advantages that such a pedagogical approach would enable. In the Connectivist model the learning community is described as a node, which is always part of a larger system. Nodes arise from the connection points found on a system. Nodes may also be libraries, websites, journals, databases, videos, or any other sources of information (Siemens and Conole, 2011). For example, additional digital content can be easily added to the VR software, including video footage, quizzes, links to websites, and so forth, to further enhance the knowledge gain.

Finally, turning to a precise area where the pairing of CR with PLS-SEM appears to have been advantageous, discussion now moves to the immersion theme. When using this novel evaluative method, the findings confirmed that immersion is a contributing factor in improving knowledge when using VR simulation within a nursing long-term conditions unit. This finding supports those of previous researchers (Lau and Lee, 2015, Witmer and Singer, 1998, Stevens and Kincaid, 2015, Tüzün and Özdinç, 2016). In Barnett et al.'s (2016) study, VR group students reported improved relevance and "realism" when compared to the Control group. Slater and Wilbur (1997) had the same view, that this "realist experience", could be nurtured in VR environments, encouraging greater immersion and improving student learning gains. However, previous research has tended to use the terms engagement, immersion and flow interchangeably (Lazaros et al., 2018). Whilst Tcha-Tokey et al. (2018) suggested that engagement is the first step towards immersion, there have been few attempts to measure the strength and direction of the precise relationships between these concepts. By pairing CR with PLS-SEM, this thesis has evidenced a statistically significant mediation effect of the engagement to immersion pathway that produced statistically improved knowledge gain in students who learnt using VR.

7.5 Unexpected outcomes

Two main unexpected outcomes were noted. Firstly, although the engagement to immersion path coefficient was found to be significant, surprisingly other path coefficients were not found to be statistically significant, as demonstrated using a two-tailed permutation-based interval test. One possible explanation for this might be because confidence was found to fully mediate the experience to knowledge pathway, and immersion was found to partially mediate the engagement to knowledge pathway. Further mediation analysis is possible with the PLS-SEM software though falls outside the scope of this thesis.

Another unexpected finding was that the confidence gain and knowledge gain of the students did not always correlate. For example, overall confidence gain of the Experimental Group was 30% compared to an actual knowledge gain of 10%. This might mean that there is a difference between what people perceive they gain in knowledge and what they learn. This finding is problematic to explain but might not be a detrimental consequence of the VR learning simulation. However, whilst confidence is desirable, the main aims of any learning tool could arguably be improvement in learning gain and competency. Contrary to expectations, the Experimental Group participants felt less confident about their pharmacological knowledge once they had used the VR simulation. Perhaps they had begun to realise what they did not know, though this assumption is speculative and further investigation around this area is warranted.

7.6 Challenges that remain

There are several challenges to using VR simulations for large group teaching in HE. Such challenges include scalability and sustainability considerations, funding/cost, technical issues, IT capabilities, and relevance to student learning in terms of pedagogical considerations. These challenges will now be discussed.

This research has described some of the challenges that arise when using VR activities in large group HE settings. This helps us to better appreciate the issues and concerns faced by educators when using such technologies to support their teaching and learning. The research has identified and attempted to explain several problems, but it is too complex for a single explanation, despite attempts at compound explanations, i.e. using configural analysis via CR and PLS-SEM. There might be other possible causes that will need further exploration in the future. Whilst this thesis has evidenced that non-immersive VR is more efficient than normative HE learning, some challenges of sustainability and scalability remain.

One area thought to be relevant to the sustainability and scalability of VR in HE is pedagogy. It is believed that to be embedded in curricula and to be considered relevant and useful to users, VR technology must be closely aligned with pedagogical aspects. Previous research concerning the use of VR in HE has been vague about the theoretical pedagogical models used, with most either not specifying the underlying pedagogy or loosely linking the intervention to constructivism (e.g., Birt et al. 2018), as was outlined in Chapter Two (the systematic review of the literature). Other researchers were more specific about pinning down the types of pedagogy involved when VR learning occurs; for example, Huang et al. (2010) supposed that VR "contextual learning" has a unique opportunity to inject educational affordances into learning situations. It is recommended that future research continues to assess how strong pedagogical planning can promote best practice when using VR simulations in HE.

Challenges to using VR technologies to support the nursing curriculum at BU were identified and reported; however, the systematic review of the literature included in this thesis revealed that only 40% of studies reported negative findings in their research. Such negative findings are the clues to any potential barriers to the scalability and sustainability of VR technology use in HE. The negative findings in the current study, namely technological difficulties with the VR (e.g., downloading issues and issues with navigating the avatar), were also reported in the literature, for example, by Hack (2016). However, Akcayir and Akcayir (Akcayir and Akcayir, 2017) trust that these types of teething problems can and should be rectified in order to continue to assess and refine teaching and learning with VR.

Software concerns were raised by both LTs and students and were the main apprehension about the VR simulation. For example, students and LTs were worried that the software could make the computer “crash”. The software did “crash” on occasions, particularly when being run on an Apple Mac laptop. Since this experiment the software has already been upgraded and is now more stable and reliable. LTs and students had also been concerned about moving the avatar around the virtual wardroom, finding that the avatar sometimes got stuck behind various pieces of furniture or equipment. LTs cited more challenges than the students did. This finding was to be expected, because LTs are responsible for helping to ensure the smooth running of the technology during a teaching session. LTs contributed to all stages of the experiment, including the design stage, piloting and iteration of the VR simulation. They will continue to be involved when the immersive version of the VR simulation is tested. Input from the LTs during this project has been invaluable and has contributed its success. It is also important to note that the VR simulation used for the purposes of this research was a proof of concept pilot version of the simulation. However, it is necessary that low-cost and scalability issues are borne in mind for future iterations.

Understandably the academic was less concerned about how the technology would work but more concerned about investment in such use of VR simulation to enhance the nursing curriculum at BU. The academic had had previous experience of using similar technology at another university and was concerned that such technologies might come and go, depending upon whether the managers thought they were worth investing in or not. She was also concerned that teaching staff might not have time to plan and develop such approaches within their present workloads. Her concerns are echoed by other academics anecdotally, and in the literature, for example, (Dolphin and O'Connor, 2019), and for this reason best practice and successful case studies are shared amongst staff. For example, the current study was shared as a Fusion Case study on the BU website (Singleton et al., 2019). Via using case studies shared between academics within the university and beyond, it is hoped that others can learn from this pilot study.

In relation to the impact of the cost of VR technology on its scalability and sustainability, the notion of fidelity is also relevant. Returning to the observation by previous researchers (Chiniara et al., 2019) that design and human affordances are more important than fidelity when evaluating the success of simulation pedagogy, the results of this thesis certainly echo this point. The software tested in this pilot study was deemed to be of low fidelity due to its low-cost nature; despite this, students found themselves immersed in the virtual ward with the diabetic patient. This indicates that future iterations of such VR simulations do not necessarily need to be of high cost and high fidelity to be successful in improving student learning outcomes and enjoyment in their learning.

Finally, investment is a major driving factor that must be considered. Findings from the qualitative results have heightened the need for initial investment in IT services support for such VR innovations, in addition to ongoing maintenance and support for the software, in order that LTs feel confident that the VR simulation will perform and respond in a reliable and consistent manner and in a variety of HE settings (e.g., the BU Yeovil site where the Internet capabilities are not as strong as those at the main BU sites). At present, companies supplying VR services tend to provide individual student subscriptions. This means that ongoing payments to the software companies are necessary for the long-term access to and availability of the software for student and academic use. It is therefore, recommended that academics wishing to use similar VR software plan this ongoing subscription into their budgeting and bid writing, or find a creative way to commercialise the software, therefore ensuring a constant supply of funding is pumped back into the HE setting in order to support ongoing costs. Cost implications and funding to support such VR innovations in the education sector remain an ongoing challenge to all involved.

7.7 Chapter Summary

The chapter was driven by the main aim of the thesis, which was to assess the effectiveness of CR paired with PLS-SEM as a method to evaluate the impact of VR simulations on undergraduate nurse education. Having recapped the main findings of the thesis, they were distilled in relation to the objectives via three areas: what has been learnt from creating the evaluative model, what has been learnt from the application of the model to VR and what did that application tell us about the innovative pairing of CR with PLS-SEM. Unexpected outcomes and remaining challenges were also deliberated. The next chapter will conclude the thesis.

Chapter 8 Conclusion, Limitations and Recommendations

8.1 Synopsis

As Harris S.J. (1917-1986) asserted: “the whole purpose of education is to turn mirrors into windows”. Therefore, in this final chapter the research will be closed by drawing conclusions about the efficacy of VR learning for nurse education. Not only does this chapter seek to reflect upon the learning of all those involved in the study, but also reflection looking beyond the confines of the lecture theatres at BU, out through the window and into the bigger world beyond. With that in mind, final consideration will be given to the potential impact of this work, particularly for HE researchers and educators.

A brief recap of the methods used will be provided prior to discussing the significance of the findings, the contribution to knowledge evidenced, the limitations of the research and recommendations for the future. A brief personal reflection on the research journey of the author is included prior to the close of the chapter via a final summary.

8.2 Rationale and Methods

The topic was chosen because student nurses required a more engaging and effective method of teaching to understand complex diabetes physiology and pharmacology, and hypoglycaemia knowledge. The gap in student nursing knowledge was identified by senior nurses on a local hospital ward who had found that student nursing knowledge about hypoglycaemia was limited. The research evaluated, via a new method, whether the use of a non-immersive VR learning simulation would be more effective at improving short-term knowledge gain when compared to the teaching method usually used within the student’s Long-Term Health Conditions Unit.

The integrated systematic review (Chapter Two) found that out of the 50 articles analysed, experimental design was only employed in 21% and just 9% had randomised participants into a Control or Experimental Group. This implies that the current study was more rigorous than most previous studies. Unlike most of the studies described in the systematic review, this study employed pre-testing to analyse prior knowledge and establish group equivalence before conducting an achievement comparison. The present study reported negative findings as well as positive outcomes. In the systematic review, less than three quarters (72%) reported their findings in this balanced way. The systematic review recommended that future studies should adhere to a “stronger evaluation” approach. Hence this study used a medium to large heterogeneous sample size, and was evaluated by students, LTs and an academic. The sample was heterogeneous in that it comprised both Mental Health and Adult Student Nurses in two different locations (Yeovil and Bournemouth).

The RCT methodology was further strengthened through the application of CR. Bhakar (1989) asserted that applying the cause and effect relationships outside of the laboratory involves moving from a closed system, to open systems where mechanisms become more contingent and indeterminate (Walsh and Evans, 2014). This is because of the convolution of open systems, requiring a more exhaustive examination in the search for causative agents. Through FGDs, analysed via a CR approach, this study considered students’, academic and LTs’ points of view. Considering the experiences of both the student and tutors was also recommended by Steils et al. (2015). Merchant et al. (2014) had highlighted this limitation in previous studies which had only involved evaluation by the user, therefore this study could be said to have a more rigorous evaluative process.

Evidence-based practice, which also includes in-depth understanding of how an intervention works in order to improve attainment, is being increasingly demanded (Pawson and Tilley, 1997b); hence a CR approach was employed for the research reported in this thesis. This CR approach aimed to respond to the complexity of social phenomena holistically (Allana and Clark, 2018). The approach is viewed as being suitable for evaluating interventions which comprise several interacting components affected by contextual elements. Education has to be regarded as an open system, for pedagogical and ethical reasons (Wrigley, 2019).

CR bridges the void between positivist and interpretivist research approaches by conjoining constructionist epistemology with realist ontology. It is particularly appropriate for mixed methods research, providing veracity and consistency in responding to multiple RQs by using a diversity of methods whilst avoiding methodological incommensurability (Walsh and Evans, 2014). It was hoped that using a CR perspective would have the potential to assist non-reductionist analysis of the fundamental associations between learning environments, educational knowledge, and the interior world of the student. Adopting a CR stance in mixed methods research offers a sound ontological basis, which supports and justifies the use of diverse methods exploring the same phenomenon, philosophically (Zachariadis et al., 2013). Thus, this study is guided by CR, providing insight into why the VR intervention would work for some people and not for others, drawing from quantitative, qualitative, and mixed methods, to explain why things happen the way they do.

This is the first PLS-SEM analysed CR study as far as is known. The PLS-SEM approach in addition to the qualitative CR element enabled a more precise attempt to find out how the VR simulation worked and how it can be improved. PLS-SEM was selected because it can evaluate relationships between mechanisms involved, not just in isolation. This has proven to be a particularly useful approach for evaluating a more complex phenomenon, in this case how a VR simulation works in a real life large-scale lecture scenario, with over 100 students (using their own varying devices) at any given time, supported by a team of three LTs and one academic. There are many mechanisms at play in this instance, with varying relationships (and strength and direction of those relationships) and differing interactions. Despite these numerous and interconnecting variables, through evaluating the data via assimilated CR/PLS-SEM, it has been possible to draw conclusions that have some generalisability to other student groups and settings.

8.3 Responding to the aim and objectives of this thesis

The main aim of this thesis was to assess the effectiveness of CR paired with PLS-SEM as a method to evaluate the impact of VR simulations on undergraduate nurse education. The aim has been achieved through two objectives, which were to determine: the effect of pairing CR with PLS-SEM as an evaluative method; and how using this novel evaluative method can inform our understanding of the impact and future use of VR simulations for undergraduate nurse education. The section will now proceed structured by these two objectives.

8.4 What is the effect of pairing CR with PLS-SEM as an evaluative method?

To conclude, the assimilation of CR with PLS-SEM as an evaluative method was straightforward, comprehensive and scalable using SmartPLS 3.0 software and step by step guidelines produced by Hair et al. (2016). The innovative method enabled the following:

- clear understanding and modelling of the abstract statistical concepts
- the application of PLS-SEM overcame the challenges of applying statistical analysis to the complex configurational approach to causation advocated by CR

- a strong visual approach that would be useful and inclusive for researchers, particularly those with specific learning needs

- richer insights into the relationships between CR themes and concerning who VR is useful for, when, where and in what circumstances

- the finding that students valued the personalised, inclusive and student-integrated aspects of the VR simulation

- identification of the need for future iterations of the VR software to take on a Connectivist approach (Siemens and Conole, 2011), by adding digital content including video footage, quizzes, links to websites, and so forth, to further enhance the knowledge gain

- stronger evidence that there is “measurable” learning gain when VR simulation is used in HE

- evidence of a statistically significant mediation effect of the engagement to immersion pathway that produced statistically improved knowledge gain in students who learnt using VR.

8.5 How can using this novel evaluative method inform our understanding of the impact and future use of VR simulations for undergraduate nurse education?

The novel pairing of CR with PLS-SEM produced rich and vast data; hence this section starts with a recap of the main findings displayed as heat maps to aid discussion. A summary of the main conclusions will then be reported by drawing upon all four data sets produced by the innovative evaluative method.

A heat map diagram was created (Figure 8-1) to compare the hypoglycaemia pre-test quiz results with those of the Control Group (paper-based learning) and the Experimental Group (VR-based learning). The heat map evidences that, overall, the Experimental Group performed better. The heat map demonstrates specific areas where the VR group gained a higher short-term knowledge gain than the other two groups. This information would be of use to educators using this specific VR simulation. It would also guide VR simulation designers to plan and create the content of their simulation realising that short-term learning gain is not achieved in a consistent manner. In other words, the VR activity is better at teaching some areas of the curriculum than others.

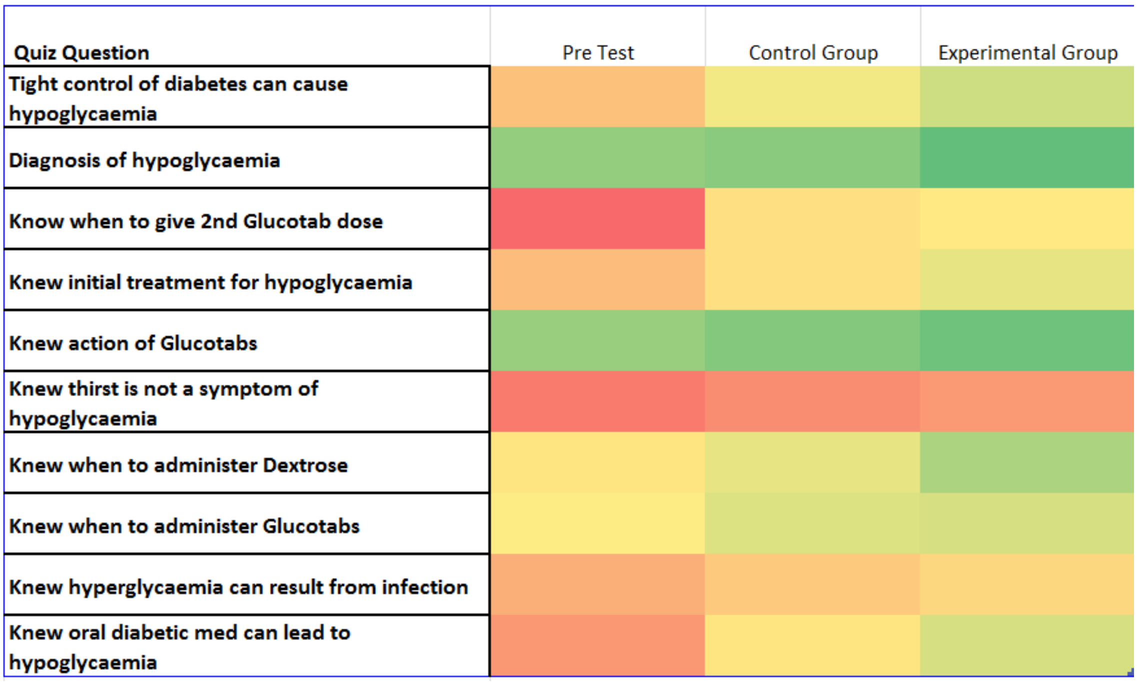


Figure 8-1 Heat Map Comparison of the Hypoglycaemia Knowledge Quiz Survey Data (Green is a higher score whilst red is a lower score).

Causal (according to the PLS-SEM definition of causal) relationships between the five variables tested, are demonstrated in Figure 8-2. As can be seen, across the data set prior knowledge (e.g., of using computing software) did not impact upon short-term knowledge gain after using either the control or experimental interventions. These results show that the VR simulation design is an inclusive learning tool, regardless of gender, age and prior computing experience.

Variable pathway	Coefficient
Experience to Confidence	
Confidence to Knowledge	
Experience to Knowledge	
Engagement to Immersion	
Immersion to Knowledge	
Engagement to Knowledge	

Figure 8-2 Relationships between variables across the entire data set

Conclusions can also be drawn in relation to the group comparison data (Figure 8-3). One unexpected outcome is that the Experimental Group did not have higher confidence scores. It is proposed that perhaps, having completed the VR simulation they had become more aware of their knowledge and skill limitations in relation to Diabetes Nursing care. The heat map shows the engagement to immersion pathway in green (Original Difference of d=*.278*). The causal pathway between the two variables was stronger in the VR (Experimental) Group.

VARIABLE PATHWAY	*STATISTICAL SIGNIFICANCE*
EXPERIENCE-CONFIDENCE	*-0.053*
CONFIDENCE-KNOWLEDGE	*-0.124*
EXPERIENCE-KNOWLEDGE	*-0.048*
ENGAGEMENT-IMMERSION	*0.278*
IMMERSION-KNOWLEDGE	*0.023*
ENGAGEMENT-KNOWLEDGE	*0.005*

Figure 8-3 Control and Experimental Groups comparison using a two-tailed 95% permutation-based confidence. If the original difference (d) of the group-specific path coefficient estimates does not fall into the confidence interval, it is statistically significant. The path coefficient concerning the ENGAGE→IMMERSION has an original difference of 0.278; as this does not fall between the confidence intervals of -0.255 and 0.262, it is statistically significant.

Through the assimilation of CR with PLS-SEM it can be concluded that the engagement to immersion pathway was the key point of action of the VR simulation. This specific part of the pathway model is thought to have been most likely to be responsible for leading to higher short-term knowledge scores in the VR group. Overall, non-immersive virtual reality was found to be significantly better in terms of short-term learning gain of hypoglycaemia knowledge than normative learning methods used for undergraduate student nurses.

The CR/PLS-SEM analysis revealed that VR learning (in this context) contained multiple mechanisms. This approach to analysis enabled scrutiny of the combinations of the mechanisms involved and their exact interactional relationships. The first CR configuration (CPMAO#1) was not supported by the research findings (which evidenced the inclusive nature of the VR simulation). However, the second CR configuration (CPMAO#2), which tested the associations and interactions between participant engagement levels with the VR simulation, affordances of VR (in particular the visualisation of abstract concepts), participant immersion with the VR simulation and finally, student enjoyment of the learning, was supported by both the quantitative and qualitative findings.

The evaluative method used in this thesis involved analysing four data sets. The systematic literature review (Data Set One), CR literature review (Data Set Two), quantitative data (Data Set Three) and qualitative data (Data Set Four) are the four data sets. In the next section, therefore, an attempt is made to pull together what has been learned from all four data sets.

Data Set One, evidenced that large-scale, experimental studies were required to evaluate VR technologies more holistically. Hence, a larger scale (n=171) experimental study was designed and implemented to evaluate the use of VR with undergraduate nursing students.

Data Set Two defined the five themes, i.e. confidence, experience, engagement, immersion and knowledge (which also shaped the conceptual framework evidenced via Data Sets Three and Four). CPMAO#2 was found to be significant and included: a contextual mechanism of participant engagement level with the learning content, programme mechanisms including affordances of VR, for example, visualisation of abstract concepts, the identified agency was deemed to be participant immersion with the desktop VR, and finally, the outcome was enjoyment of the learning activity.

From Data Set Three, a striking result is that having completed the case study the Experimental Group had higher mean values for engagement, confidence and knowledge. The point of action for the pathway model appears to be the engagement to immersion pathway, which was found to be statistically significant (as measured by the permutation test).

Finally, Data Set Four demonstrated that students found that the VR Exercise aids understanding of the complex concepts associated with hypoglycaemia, provides immediate feedback about their clinical decisions, can be completed multiple times, for example, for revision/distance learning, aids visual learners, complements ward and simulated ward experiences, and provides more opportunities for safe practice.

Drawing upon all four data sets used in the pairing of CR with PLS-SEM, the main conclusion is that despite acknowledging at least two main challenges (namely software instability and avatar navigation issues) of highly interactive VR simulation software, this research claims that non-immersive VR is more efficient than normative HE learning. This claim is based upon the unique CR/PLS-SEM analysis that demonstrated statistically significantly higher short-term knowledge gain in the Experimental Group, resulting from an increase in students' confidence, stronger student engagement, and consequently fuller immersion in their learning. This is further supported by the findings from the qualitative analysis which revealed that superior engagement in learning was perceived to be the main advantage of VR simulation.

8.6 The significance of the findings

The most significant finding is that the pairing of CR with PLS-SEM was comprehensive and scalable for the user. Moreover, it permitted valid and reliable modelling of the statistics, whilst providing a visual approach that would be useful and inclusive for researchers. Through using this innovative method, a more in-depth understanding of relationships between themes was gained and a strong sense of student voice permeated the findings (e.g., the appreciation of the personalised aspects of VR simulations).

Another significant finding to emerge from this thesis is that whilst multiple mechanisms interact in different ways for different learners, engagement leading to immersion is a key mechanism when learning using VR, as has been evidenced from both the qualitative and quantitative findings of this research. This conclusion will be of interest to software designers and educators in encouraging them to ensure that activities they create, and use are both engaging and immersive in order to produce the best outcomes in terms of student learning. In this instance the immersive quality of the VR simulation was thought to be attributed to drawing upon the knowledge of Diabetic Nurse Specialists from the local hospital, a lecturer who usually teaches the unit, and through designing and modelling props from photographs of the nursing equipment used on the local hospital wards. These aspects led to students feeling as if they were looking through their “ward placement” eyes and thinking and acting as if they were on a hospital ward. The present study has gone some way towards enhancing our understanding of the specific mechanisms of action that interact when students learn via VR.

Another significant aspect of this research is that improved knowledge was evidenced for the Experimental Group via statistical modelling. Large-scale experimental design analysed via the robust approach (PLS-SEM) provides confidence in the results and permits the claim that the diabetes VR simulation was effective in improving student learning when compared to normative instruction for 2nd year Adult and Mental Health nurses.

8.7 Contribution to knowledge

In terms of implications for scholarship and the literature, this research is original methodologically, as it is the first of its kind to pair a RCT, analysed via PLS-SEM, with a CR angle for evaluating a VR intervention in HE. It builds upon accepted findings and theories of other researchers and then extends beyond previous knowledge by incorporating original elements.

There are five main areas in which an original contribution to knowledge is evidenced:

1. The pairing of PLS-SEM with a CR review of an RCT
2. Two reviews of the literature (integrated systematic and CR reviews)
3. Development and testing of an original conceptual model
4. Application of non-immersive VR to large group teaching in a UK HE setting
5. Collaboration by LTs, academics, software developers, and clinical staff at all stages of the research (and students at the final stages).

This thesis has been one of the first attempts to thoroughly examine the use of VW, VR, and AR use across the HE sector. The review concluded that large-scale, experimental studies are required to evaluate these technologies more holistically; this was then acted upon in the empirical work. Literature drawn together through the CR review provided evidence to support the construction of the conceptual pathway model. The contribution of knowledge was to assimilate the findings of these reviews and to test the conceptual theories through conducting empirical research.

The Conceptual Model proposed (containing some original pathways) tested whether significant “causal” relationships existed between identified latent variables. This combination of findings provides confirmation for the conceptual premise that when using a VR simulation (of the type evaluated within this thesis) for the teaching of clinical skills in an HE setting, engagement in learning leads to immersion and subsequently higher short-term knowledge gain. The conceptual model was analysed using a unique combination of CR evaluation of an RCT and PLS-SEM.

Past research has been exemplified by small scale bespoke VR activities (e.g., collaborative VW activities). Chapter Two reported the activities that have been most often used when teaching with VW/VR/AR in HE. The systematic review revealed that VW/VR/AR technology was only used in a lecture theatre setting in 5% of instances. By contrast, this thesis reports the findings of the use of non-immersive VR within larger group (100-200 students) lecture theatre settings.

Walsh and Evans (2014) advocated a partial use of knowledge transfer procedures where VR makers are tasked not just with manufacturing VR experiences but also engaging with educational providers and users to understand their needs and bring that to the design process. This thesis reports a process that connected students, researchers, academics, clinical professionals, LTs and software developers in order to evaluate the VR technology via a holistic CR approach. This approach enriched the research and provided various viewpoints.

In relation to implications for policy and practice, by reviewing the recent studies on the use of VW/VR/AR in HE and adding new findings, it is hoped that this study can provide policymakers and practitioners in the education sector with new insights into effective methods to evaluate the application of these technologies. The findings of this research also provide insights for people designing VR software and developing curricula. For example, the results of this research have indicated that educators and software developers should design VR simulation exercises with engagement and immersion in mind. The results are supported by previous researchers who found that identifying means to encourage and develop presence needs to be part of the learning and teaching strategy and design of Virtual Environments.

Overall, this study strengthens the idea that prior experience (age and computing) is not essential when learning with VR. The VR simulation created and tested was found to be an “inclusive” tool for teaching and learning. Moreover, this research has demonstrated that VR simulations (such as the one created and tested in this project) provide an opportunity for experiential learning and safe practice of clinical skills.

It is also argued that the VR exercise was highly interactive and encouraged personalised, situational and Connectivist learning (though more opportunities for Connectivism can be added in future iterations). This is due to the unique affordances of VR and means that students’ learning progressed through Kolb’s learning cycle (Baker et al., 2012). The instant feedback enabled via the VR simulation, was deemed to be a clear advantage in accelerating student learning of the concepts involved in diagnosing and treating a deteriorating patient suffering from hypoglycaemia.

Previous researchers (Inman et al., 2011), stressed that there should be a clear connection between course learning objectives and the VR exercise, and that the training content should be adapted to the curriculum. This advice was implemented at the first stage (the design stage) of the research and should help to ensure that the VR simulation continues to be used in the coming years, as it promotes sustainability and scalability of the teaching innovation.

The findings of this study have several important implications for future practice. Despite the strengths and advantages, three main challenges have emerged from all four data sets, namely:

- investment in the use of VR
- technical issues including stability of the software and downloading the software
- student familiarity with the software.

In terms of investment, both money and time were found to be equally important considerations. It was highlighted that it can take a varying amount of time for students to download the software. This investment in time will be worth it as more scenarios become available. The time problem might also be overcome in the future through the use of an application version of the software that can be used on a mobile phone or tablet device (though it is also acknowledged that accessing the software in an "immersive" route, e.g., via Google Cardboard headsets, might add additional drawbacks such as causing nausea or dizziness to the user).

Academics and LTs involved in this research expressed concerns about the lack of "investment" in these VR technologies from managers. They felt that managers did not appreciate the additional time required to evaluate, create, and implement new technology innovations such as VR, and the curriculum in which such technology use should be embedded. Monetary investment was also thought to be lacking to support such innovations within an educational setting.

Technical issues relating to the downloading and stability of software within an HE setting were cited by all involved in this study. A reasonable approach to tackle this issue could be to move to a point where appropriate systems, services, and support for VR become a priority for HE IT services and faculties at BU. However, there would be various steps necessary to achieve this. Hence further investigation to unpick these steps is required.

Student familiarity with software was the main challenge associated with this research. The systematic review (Chapter Two) led to the recommendation that when using VR technology, students required the relevant skills needed to accomplish the activity. Students need time to practise moving the avatar around and making selections in such highly interactive situations as the VR simulation created and tested in this research. In this study the learning was flipped to allow time within the curriculum for the students to complete the VR exercise. Students had been instructed to download the software and run through the tutorials at home prior to the seminar. However, there is a definite need for planning and allowing time within the usual teaching time for students to complete tutorials and receive support from LTs and academics. Once again, this would be worth the time invested once a series of VR activities has been created.

As with other dynamic media before virtual reality, VR has been available before being well understood (Steils et al., 2015). However, this research has contributed to pedagogic knowledge related to the implementation of VR exercises within traditional lectures from students', academics' and LTs' points of view. The intervention results should influence pedagogical practices and methods in education in general, and a dissemination pathway with practitioners will be designed to this effect. It is expected that the model generated by this project can be extended to improve the end user experience of VR technology users beyond the lecture theatre, with transferability to different learning contexts, for example, nursing homes, hospitals, and the community.

During the Covid-19 pandemic, education providers used a range of online platforms to deliver live and pre-recorded lectures and recreate seminars. Uncertainty around when students and staff would be safe to return to face to face learning resulted in a continuation of distance learning approaches. Nurse educators planned to bring back student nurses in reduced groups to carry out clinical skills and practice elements of the courses from the autumn 2020. However, this introduces issues such as the need for flexible timetabling and requires staffing by educators who have not been categorised as being in a vulnerable group. It also necessitates the willingness (and ability) of students to return to campus whilst the disease is an ongoing threat. The type of VR simulation evaluated and discussed in this thesis would be beneficial to further enrich student's online learning experiences, particularly (though not exclusively) during times of National and International pandemic.

Nurse educators who worked on this VR simulation project have extended the possibilities of what can be done in the physical classroom- beyond the boundaries of the VLE. A more authentic, work-based theory-practice bridge was created that in future can provide greater insight for students who are in remote classrooms.

8.8 Limitations

One limitation of this study is that only non-immersive VR (accessed via desktop computer) was tested. Several other researchers have found that immersive VR is superior to non-immersive VR in training students. For example, Makransky et al. (2019) concluded that immersive VR is superior to desktop VR in arousing, engaging, and motivating students. Future iterations of the software will be accessible via mobile phones.

Other elements that could have been included are longitudinal data and video recordings of the VR sessions. The latter was explored with an LT and was then rejected as it was thought that within a large lecture theatre setting it would not be possible to capture individual participant facial expressions. Whilst this study did not confirm long-term knowledge gain, in relation to diabetic hypoglycaemia, it did substantiate a short-term knowledge gain. Longer-term knowledge retention was not possible to confirm because the study was not of a longitudinal design. The results therefore need to be interpreted with caution.

Psychological factors affecting immersion need to be deliberated in relation to learning outcomes (Michailidis et al., 2018). This thesis argues that other factors, which are beyond the parameters of this study, might also influence immersion/presence and learning outcomes. These factors might include philosophical thought, cultural beliefs, spirituality, mental health, imagination, and states of consciousness and wakefulness. Neurophysiology could also have an effect, for example in the case of ASDs, and ADHD a person’s “filter” can be different in terms of how they perceive incoming sensations. Finally, drugs/medication, autosuggestions, hallucination and sensory problems, for example, macular degeneration, loss of pain sensation, and loss of smell, might all contribute to a person’s level of immersion or how they perceive what is going on around them in a virtual environment. These factors have not been included as part of the research reported in this thesis, because they are beyond the scope of the aims and objectives of the study, though it is acknowledged that they might play some part in learning via VR.

In relation to validity and reliability, for the systematic review of VW/VR/AR use in HE, Kitchenham’s (2004) guidelines for conducting systematic reviews were adhered to. This means that the review is transparent, comprehendible and can be replicated by others. EPPI Reviewer software was also employed for the early stages of the systematic review, adding to the rigour and reliability of the process. A small team (n=4) of librarians and academics co-authored the review protocol which included generating and testing the search string. Having produced a review protocol, a second reviewer (LF) was employed to conduct joint reviewing of 100 journal articles to ensure that the first reviewer's procedures and practices adhered to the systematic review protocol.

Widely used and evaluated survey procedures and guidelines were selected for this research. For example, the participant surveys were created using Witmer and Singer's (1998) Immersive Tendencies Questionnaire. Analysis of the qualitative data followed thematic analysis widely used to draw commonalities from the data. For the CR review the guidelines of Lee et al. (2010) were followed. Using these tried and tested methods improves the validity and reliability of the procedures conducted in this research. However, some limitations exist in relation to the evaluation of the VR experience. For example, students rated their own confidence using Likert scales. To improve the rigour of this evaluation in future research, a Confidence scale could be used. One such relevant scale is the Nursing Anxiety and Self-confidence with Clinical Decision-making Tool (Zieber and Sedgewick, 2018).

Reliability tests try to identify the consistency of the measuring instrument whereas validity tests try to find out how accurate an instrument measures a particular concept it is designed to measure (Rajasekar et al., 2017). In relation to validity, both the internal and external validity of the study were considered.

Internal validity is an evaluation's precision in identifying the causal effects of the intervention – i.e. whether alterations in the outcomes observed resulting from the intervention are the result of the intervention, rather than participants' features. One way of improving an evaluation's internal validity is through the tight control of aspects, including the selection of participants, assignment to intervention conditions and implementation of the intervention (Fredericks et al., 2019). Reeves et al. (2018) stated that this is best achieved with an RCT, which is considered to be best practice for inferring the causal effect of interventions. Hence, internal validity was optimised in this study.

Conversely, external validity denotes the ability to reproduce the intervention's effects in different contexts and with different students who will vary in their characteristics (Robinson et al., 2016). External validity requires minimal control of various aspects of the evaluation, to ensure participants and intervention conditions represent those encountered in other HE situations (Sidani et al., 2017). This research aimed for a pragmatic approach to research that balanced internal and external validity, and generated evidence that is relevant to HE institutions.

Beyond the sampling quality, and the question of validity of the measurement instrument in relation to individuals beyond the actual sample used for this study, Hair et al.'s. (2016) guidelines were adopted for the validity and reliability testing. Evaluation of the measurement model included loadings, Cronbach's alpha/composite reliability/Ave, HTMT (for reflective variables) and redundancy analysis, VIF, and significance and relevance of the indicator weights (for formative variables).

Evaluation of the structural model included examination using collinearity assessment, coefficients of determination (R^2) (and f^2 effect sizes), predictive relevance (Q^2), and the statistical significance and relevance of path coefficients (Hair et al., 2016). The Q^2 measure is an indicator of the model's out-of-sample predictive power or predictive relevance. When a PLS path model exhibits predictive relevance, it accurately predicts data not used in the model estimation. In the structural model, Q^2 values larger than zero for a specific reflective endogenous latent variable indicate the path model's predictive relevance for a dependent construct (Henseler and Hubona, 2016, Sinkovics et al., 2016, Karkar et al., 2019). It is accepted that Q^2 values larger than 0 suggest that the model has predictive relevance for a certain endogenous construct.

For this study, it was established that the PLS-SEM structural and measurement models were above thresholds for validity and reliability, according to Hair et al.'s (2016) guidelines. For example, high levels of internal consistency reliability, convergent validity, indicator reliability and discriminant validity (via HTMT) were supported. This means that interpretations taken from the results of the PLS-SEM model can be deemed to be accurate. However, since the analysis phase of the study Hair has published 2019 guidelines that include additional validity and reliability measures. These new more rigorous measures could be utilised for future research. The pairing of CR with PLS-SEM is the first of its kind and therefore requires further testing. It is planned to test the method further by applying it to the evaluation of music therapy for palliative care.

The generalisability of these results is subject to certain limitations. For instance, the scope of this study was limited in terms of the technology used; namely, only desktop VR was tested. Several other researchers (Makransky et al., 2019) have found that immersive VR is superior to desktop VR in training students. For example, they concluded that immersive VR is superior to desktop VR in arousing, engaging, and motivating students. However, there is no overwhelmingly conclusive evidence that immersive systems are more effective in educational applications than their non-immersive (e.g., desktop) counterparts (Ogbuanya and Onele, 2018). Hence, further research is indicated.

Despite its limitations, the study certainly adds to our understanding of VR use in HE settings, particularly for undergraduate nursing students using deteriorating patient case studies. It is reasonable to imply that other similar student nursing cohorts would benefit from the technology created and tested through this research. Despite its limitations, it is believed that this research has generated a new technological intervention, which can be adapted beyond the HE context.

8.9 Recommendations

Within this section recommendations are set out in relation to future design and implementation of VR technologies for nurse education and HE in general. Recommendations are also provided for future research. It has been evidenced that to further improve the efficacy of VR technologies in HE, participating students require a structured orientation or pre-briefing before the session, which includes the learning objectives, structure and process of the simulation, and familiarity with the simulation environment, equipment, manikin, and monitoring devices. Such pre-briefings help to reduce student anxiety.

Several students and one LT recommended that providing students with a choice of how to navigate the avatar around the virtual space would be a useful enhancement to the exercise. This once again would take into consideration differences in the way students learn and work. With the headset version in mind, moving the avatar would be simplified and would operate as a simple point and click teleportation system. However, the laptop access version would still be made available in future iterations so that students could choose how they would access the software. This is imperative for several reasons; firstly, some users experience nausea or headaches when using VR headsets and it is also difficult to use such headsets if the user has certain visual impairment, for example. Secondly, some users expressed a dislike of VR headsets; the reasons behind this are diverse but can include feeling self-conscious. Finally, not all students possess a mobile phone of the specification necessary to run the software. In response to this issue, BU is considering purchasing several Oculus Quest™ headsets. These headsets are standalone pieces of equipment that do not need to be attached to an additional computing device. They are also reasonably priced and thus lend themselves to small group work with students. However, for students to be able to practise the VR simulation at home they need to access the software on their own device (ideally), hence the laptop and mobile phone versions will be made available in the future.

In addition to comments made about how to move the avatar, other enhancements to the software were recommended. For example, the additions of a Chatbot feature. This had not been incorporated into the proof of concept iteration of the VR simulation in order to keep costs low, though it is recognised that the use of Chatbot would enable further learning gain for some students, particularly those who learn via listening or those who learn with a combination of looking and listening. It is recommended that future iterations of the VR deteriorating patient simulation incorporate Chatbot features in order to increase the sense of immersion and presence for the learners (Yi Fei et al., 2017) and to further support students with ALNs (e.g., dyslexia).

The issue of when in the Unit of work the VR simulation would feature was discussed, with some participants suggesting using the simulation prior to their end of Unit exam, whilst others felt it would fit agreeably between the lecture and seminar. Taking on board all the comments relating to how the simulation could help bridge the theory-practice gap, it would be well placed between the teaching Unit and ward placement. Owing to the fact the VR simulation is digital, remote and can be used anywhere at any time (unlike any AR version), the simulation could be made available for students to use at a time/s that suits them individually. This is viewed as an advantage of this approach to teaching and learning.

Overwhelmingly, students were very excited about using the VR simulation and many requested other simulations to be developed, including different scenarios (e.g., community settings) and using different avatars (e.g., patients with dementia or paediatric patients). A real opportunity was identified for using similar VR deteriorating patient simulations within mental health settings. A major challenge for mental health nursing students is to experience enough situations involving an escalation of a mental health episode, one that leads to the necessity of restraining the patient. Such high-risk volatile situations have been simulated in VR for mental health nurses to gain experience, for example by Verkuyl et al. (2020). Within clinical practice it is not always safe or appropriate (in terms of patient dignity) for students to be involved in such situations. Moreover, the number of experiences that each individual student can observe can be variable due to time factors and placement setting availability.

Taking on board the suggestions from the students, academics, LTs and software developers was part of the Design-Based Learning approach that was adopted (Hachaj and Baraniewicz, 2015). Design Based Learning includes the incorporation of feedback on the performance of the intervention into future versions. Using this process improved the likelihood of the future success of the VR simulation for all involved and the chances of it being embedded into the ongoing BU Nursing Curriculum. Embedding of an efficient teaching and learning method is viewed as superior to a “bolting on” approach because it means that the intervention is aligned to Unit learning objectives, assessment considerations as well as looking ahead to the new NMC standards of Proficiency for Registered Nurses (NMC, 2018).

The approach to creation of the VR simulation tested in this research is argued to be a real strength because it was developed through discussion with specialist nurses, nurse educators, LTs and software developers. However, one area for improvement would be that students should also be included at an earlier stage in the creation of new learning activities. During this research students were invited to provide feedback once the software had been developed. Earlier feedback from this group of participants would be advantageous.

Finally, an essential future development suggested by participants involved in the study was to include further opportunities for immediate feedback. According to Traxler (2018), assessment, including formative types i.e. “feedback” and feedforward, within UK HE, receive consistently poor levels of satisfaction from students when they complete the National Student Survey. As evidenced in the current study, feedback was associated with an enhancement in learning gain from training with the VR simulation. A major advantage of the type of VR program developed was that additional digital content can be added easily that can provide further opportunities for teaching and learning. For example, a video demonstrating the correct procedure for testing blood glucose can be included prior to the nurse avatar carrying out that nursing procedure. Students were already given immediate feedback about their clinical decision-making, along with the opportunity to retry their choice and learn about the consequences of those choices. The software has capabilities to develop this area further.

General **recommendations** to more effectively enable VR technology integration into teaching and learning in HE, in terms of **scalability and sustainability**, are that:

- Academics, clinical staff, IT services/LTs, students and software developers should work cooperatively during the planning, development, piloting, implementation, and reiterative stages of clinical VR simulation creation for use in HE

- VR simulation should not be a bolt on activity but embedded within the ongoing curriculum

- There should be training provision (and time for this training) for faculty members on the advantages and the use of VR technologies (including VW/VR/AR) in the classroom

- There should be provision for administrative, IT and LT support to reduce faculty members' workload while implementing new educational strategies

- Collaboration between faculty members to share ideas about enhancing educational strategies (discussion groups, email groups, and so forth) should be encouraged

In addition to recommendations for how, when and where VR technologies can be designed and implemented, this research has also highlighted some potential areas for future research. Despite the promising results, questions remain. There are still many unanswered questions about the use of non-immersive VR for large groups of undergraduate students. For example, the current results show a difference in learning stages, as defined by Bloom's taxonomy; further research into the other stages would be of interest. This study looked at the lower ends of the learning hierarchy – remembering and understanding. However, VR may compare differently to traditional methods for applying, analysing, evaluating and creating; this warrants further research.

Further research could also provide more evidence about "in what circumstances VR is useful". This is important because VR has the potential to assist in meeting the identified challenges associated with clinical practice experiences, such as increased student numbers and decreased clinical placement opportunities in skill acquisition. As the use of VR becomes more widespread it is important that nursing and midwifery educators take a leading role in the conceptualisation, design, integration and research of this rapidly emerging educational technology within the context of nursing and midwifery programmes. As a highly practical profession, nursing will continue to need experiential, skills-based learning, blended between the classroom and the clinical setting (O'Rourke and Caramanica, 2019). Further research could evaluate how VR technologies could be used within a skills lab setting, or indeed during hospital placements, led by practice educators.

The research has described some of the challenges that arise when using VR activities in large group HE settings. They help us better appreciate the issues and concerns faced by educators when using such technologies to support their teaching and learning. This research has identified and attempted to explain several problems, but it is too complex for a single explanation, despite attempts for compound explanations, i.e. using configural analysis via CR and PLS-SEM. There might be other possible causes that will need further exploration in the future.

Further research is also required to:

- further develop the pairing of PLS-SEM with CR evaluations of RCTs
- establish the viability for further VR scenarios
- add further feedback to the VR scenarios including videos, and Chatbot options for students with ALNs
- compare immersive VR with non-immersive VR via a large-scale longitudinal RCT (to establish long-term knowledge retention)
- test at which exact point student enjoyment acts within the proposed conceptual model
- evaluate cost implications of VR technologies
- identify best practice for building upon pedagogical foundations when developing VR simulations

8.10 A Personal Reflection upon my Research Journey

According to Zulfikar and Mujiburrahman (2018), reflection helps educators to generate feedback for the development of their classroom practices and teaching repertoire; hence this section of the thesis. As with any research, at the end of this research journey I feel that there remain more unanswered questions than I began with. However, I hope that this reflects my quest to continually develop and improve my teaching and learning strategies for the good of the students whom I teach. In this personal reflection, I will begin with the successes of my research, new ideas and techniques that I have learnt during the journey and finally challenges that I faced along the way and how some of these challenges were overcome and why and how some challenges remain at the end of that journey. I have included my field notes as part of Appendix Seven, and they go some way towards expressing the practical issues that were overcome during this research.

In terms of new learning, something that strikes me, but is of no surprise to the LTs who assisted in this research, is how many students, despite emails providing links to the surveys and software to download, and so forth, do not show up prepared for their lecture. To illustrate this point, few students attended the lecture with the software preloaded onto their device and having practised the tutorials on how to use the Fieldscapes software. It also became apparent that the Yeovil campus Internet capabilities were insufficient for students to simultaneously download software during the lecture. On top of this, one of the academics involved had used VR in a previous HE setting and relayed details of all the challenges that had been faced and how in the end the VR intervention (a SL VW activity) had been abandoned. However, talking to this academic provided a better insight into how managers of HE settings can better support and prioritise such innovations in teaching and learning.

I was beginning to feel as if the "experiment" would not be successful. However, the expression on students' faces when they were given the opportunity to use the software in class was a memorable moment. Students were excited and engaged in their learning. Even when glitches in the software arose, particularly when using Apple Mac laptops, students persevered and were able to complete the exercise. Students spotted the potential for such a VR simulation to add value to their learning and not only strengthen the theory-practice gap but provide safe opportunities to practise clinical skills and make mistakes. Despite the software being a low-cost pilot version with low fidelity, it helped students to visualise their learning. Even more promising was the overwhelmingly positive experience reported by students who had ALNs, including dyslexia, and from those with limited past computing experience. We had co-created an "inclusive" TEL tool and one that improved students' knowledge and understanding of a complex deteriorating patient case study, at least in the short-term. The research was overall a positive experience for the students.

There are of course challenges that remain unresolved. LTs were wary that the technology might “let us down” and waste students’ learning time. LTs were very aware that students were paying for their lectures and any software glitches were deemed a great failure and embarrassment in their eyes. For future iterations, LTs want to feel confident that the software will be well supported and maintained and will not let us down. For this to be realised, considerable planning, support and investment needs to be provided from the top down. As well as IT and financial support, such innovations require support from unit leaders to prioritise these innovations and provide time within the curriculum for students to engage with software tutorials, perhaps by flipping the normative learning methods.

I believe that research, including nursing research, should be useful. A finding from my research, that is important to me, is that prior to the intervention, students were unfamiliar with the Hypobox, the medication contained within it, and how to use each medication in line with Diabetes UK guidelines. Through producing the training exercise, I liaised with other local authorities and have found that Hospital Hypoglycaemia Guidelines also vary; this could be of significance when nurses move from one hospital to another. During their nurse training and Healthcare Assistant training, students are made familiar with key pieces of equipment on the ward, for example the location of the “Crash” (resuscitation trolley). This finding pinpoints a need for all staff members to become familiar with the Hypobox and the procedures and flow diagram for treatment of hypoglycaemia as part of their induction to the ward. These initial findings were presented as a poster at the Diabetes UK Professional Conference 2019, to raise awareness of the importance of familiarity of the Hypobox and how to treat patients experiencing hypoglycaemia.

I believe that through sharing these exciting insights into VR use in HE with my colleagues at BU and beyond it might be possible to reignite interest in similar innovations. I feel pedagogically inspired to investigate further, for example by comparing student experiences of the software as an app version. It will be interesting to see what can be learned about scalability and sustainability when students use the VR case study via mobile phones.

8.11 Thesis Summary

The line of argument in this thesis has been as follows. Virtual technologies have been, and continue to be, of significant interest to HE educators. There have been many research studies carried out into the efficacy and acceptability of these technologies. But, this research (via systematic literature review) found that there are significant methodological shortcomings in many of those studies, particularly with respect to understanding the mechanisms of the effect of virtual technologies on learning. Most papers are superficial and concentrate on the TAM model of usability and ease of use. Some carried out perfunctory assessments of learning effect, but predominantly by measuring student enjoyment via subjective self-reporting.

This research has identified Critical Realism coupled with PLS-SEM as a methodological approach that could enable us to understand the pathways to learning that students experience when using virtual technologies, thereby adding to our knowledge about both virtual technologies in education, and about the methodological approaches that might move the evaluation field on, past its current superficial approach.

The CR literature review identified the latent variables that form the pathways that were tested in the experiments. Therefore, two literature reviews were conducted. The experiments were in the field of health practitioner education – specifically nurse education in respect of diabetes management. Hence, the thesis is not claiming generalisability from the study, but an addition to knowledge about how the novel methodological approach taken has the potential to deepen our understanding of <u>how</u> virtual technologies might affect learning. Recommendations for policy, practice, and further research have been made on this basis.

The findings of this research complement and extend those of earlier studies and have significant implications for the understanding of how the mechanisms involved when using VR interact, and consequently how VR can be successfully designed and implemented for learning within HE contexts. This study has raised important questions about the nature of VR learning. VR may deliver greater access to practice opportunities in HE, spanning the gap between the formal and practical learning of professionals – a vital step in developing students' proficiency.

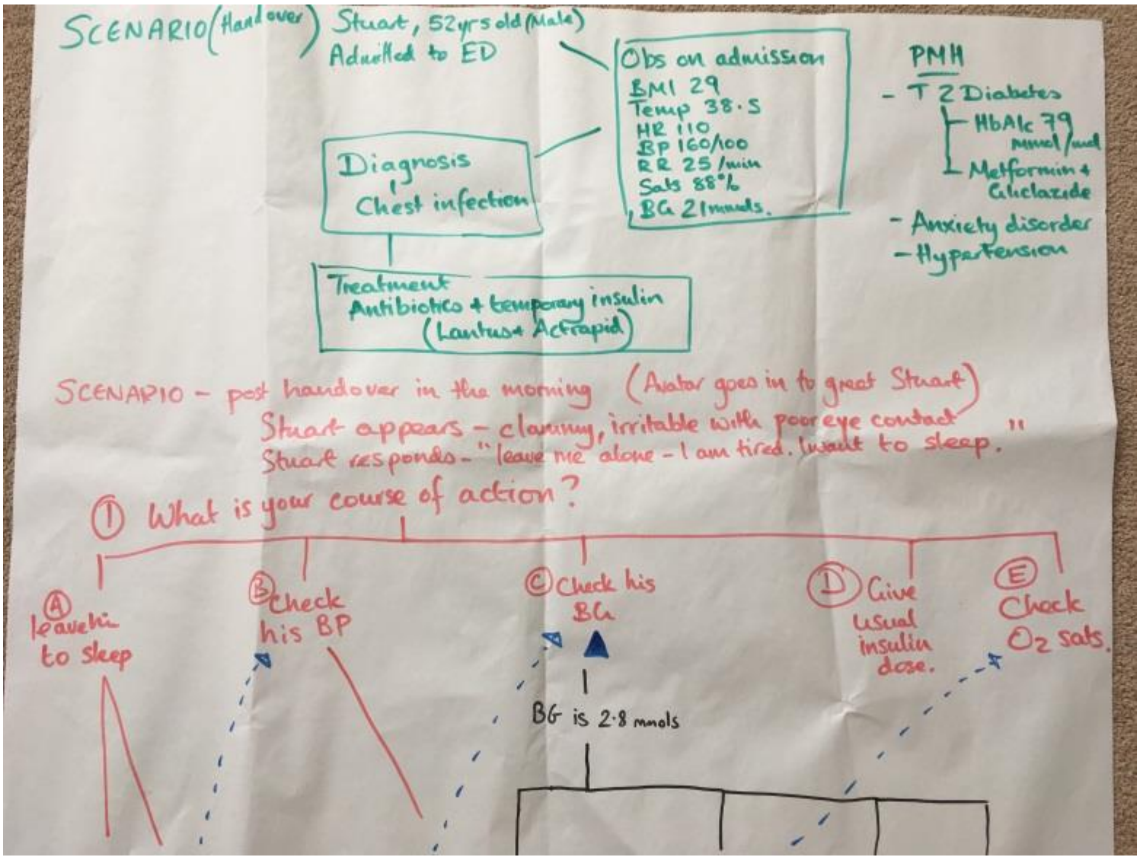
This final chapter sought to answer earlier RQs by considering data, theories, and arguments to evidence conclusions made. These concluding arguments, singly or in combination, fashioned the fundamental insights of the thesis and were the knowledge and concepts that were intended for the reader to absorb. Finally, the chapter closed with a personal reflection of the research in the first person, to capture new learning that could not be reflected in the main body of the thesis. In the words of Ruskin J. (1819-1900):

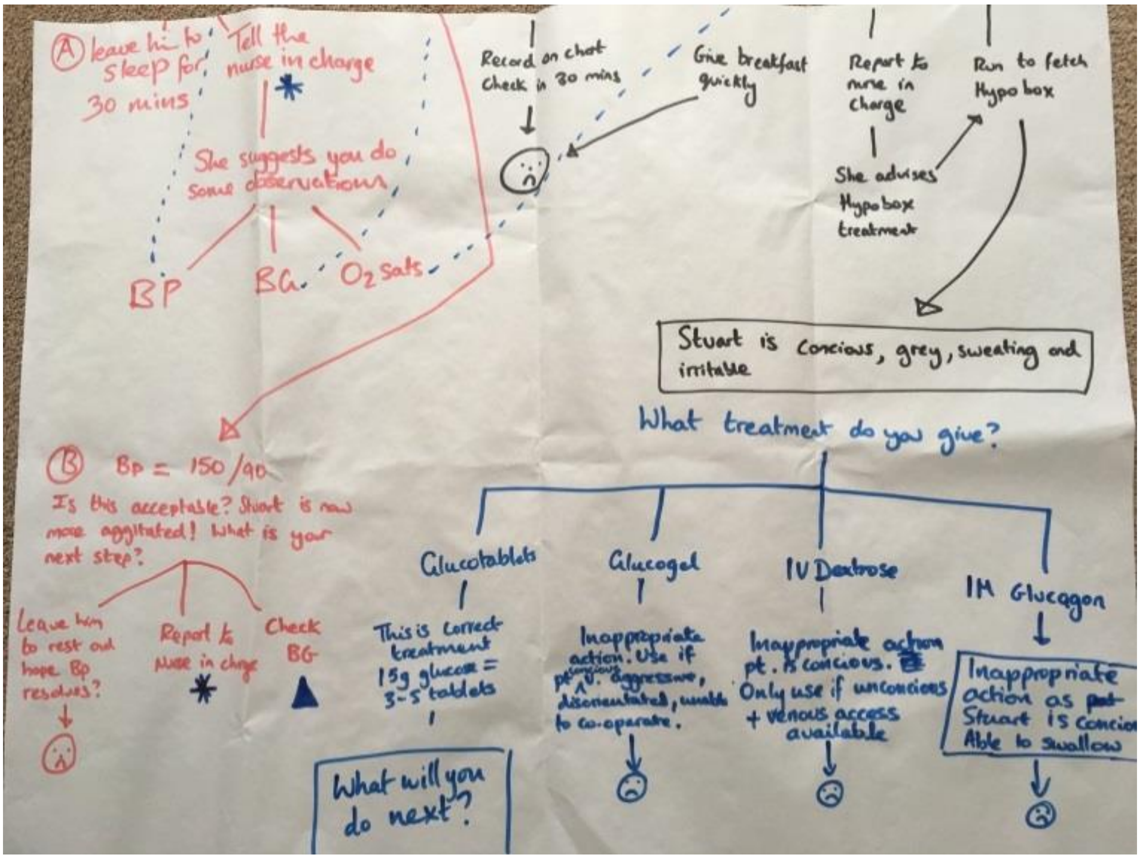
> *"Education does not mean teaching people what they do not know. It means teaching them to behave as they do not behave."*

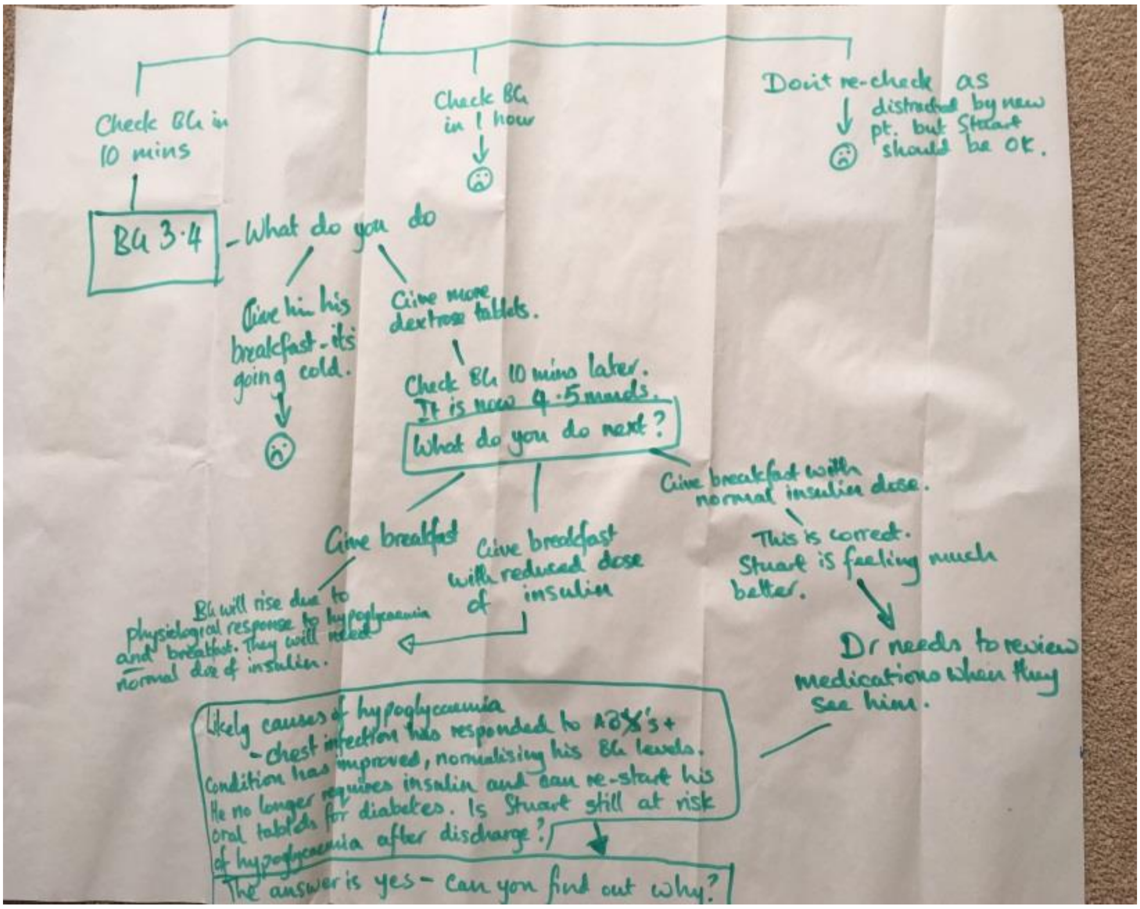
It is hoped that this thesis and subsequent publications can provoke thought amongst HE academics and encourage them to consider and explore how VR can be used in their settings to inspire student integrated learning.

Appendices

Appendix One: Tree Diagram Planning of the Deteriorating Patient Case Study




A leave him to sleep for 30 mins
Tell the nurse in charge
She suggests you do some observations
BP
BG
O2 sats
Record on chart Check in 30 mins
Give breakfast quickly
Report to nurse in charge
She advises Hypo box treatment
Run to fetch Hypo box
Stuart is concious, grey, sweating and irritable
What treatment do you give?
B BP = 150/90
Is this acceptable? Stuart is now more aggitated! What is your next step?
Leave him to rest and hope Bp resolves?
Report to nurse in charge
Check BG
Glucotablets
This is correct treatment 15g glucose = 3-5 tablets
What will you do next?
Glucogel
Inappropriate action. Use if pt. unconcious, aggressive, disorientated, unable to co-operate.
IV Dextrose
Inappropriate action pt. is concious. Only use if unconcious + venous access available
IM Glucagon
Inappropriate action as Stuart is concious Able to swallow


Check BG in 10 mins
Check BG in 1 hour
Don't re-check as distracted by new pt. but Stuart should be OK.
BG 3.4
What do you do
Give him his breakfast - its going cold.
Give more dextrose tablets.
Check BG 10 mins later. It is now 4.5 mmols.
What do you do next?
Give breakfast
Give breakfast with reduced dose of insulin
Give breakfast with normal insulin dose.
This is correct. Stuart is feeling much better.
Dr needs to review medications when they see him.
BG will rise due to physiological response to hypoglycaemia and breakfast. They will need normal dose of insulin.
Likely causes of hypoglycaemia - chest infection has responded to ABX's + condition has improved, normalising his BG levels. He no longer requires insulin and can re-start his oral tablets for diabetes. Is Stuart still at risk of hypoglycaemia after discharge?
The answer is yes - Can you find out why?

Appendix Two: Paper-based Case Study- Deteriorating Patient

You arrive at the ward to start your day shift. The Senior Nurse working with you gives you the handover notes before the patient's breakfast time.

Handover notes: 52 yr. old Stuart was admitted to ED yesterday with a chest infection. He suffers from anxiety disorder, hypertension and Type 2 Diabetes. His usual diabetic medication is Metformin and Gliclazide, but this has been stopped and he is on temporary Lantus and Actrapid. He received oxygen therapy overnight.

His observations: last evening was: Temp 38 °C, HR 92, BP 150/90, Oxygen Sats 92%, BG 16 mmol/L. When you approach him, you notice he looks clammy, pale and irritable. He says: "leave me alone, I am tired, I want to sleep!"

1. What is the best course of action?

Check his BP (Go to Q2)

Check his BG (Go to Q3)

Let him sleep (Go to Q4)

2. Stuart's BP is 150/90. Stuart's condition is now deteriorating.

What is the best course of action now?

Check his Oxygen Saturation levels? (Go to Q5)

Let him sleep (Go to Q4)

Check his BG (Go to Q3)

3. Stuart's Blood Glucose level is 2.8 mmol/L.

What is the best course of action now?

Use the Hypobox (Go to Q7)

Give him his breakfast (Go to Q6)

Let him sleep (Go to Q4)

4. Stuart sleeps for 10 mins. When he wakes up his condition has deteriorated. He still reports that he is feeling tired and unwell and is having palpitations.

What is the best course of action now?

Check his BG (Go to Q3)

Check his Oxygen saturation levels (Go to Q5)

5. Stuart’s Oxygen saturation levels are now 95%.

What is the best course of action now?

Check his BG (Go to Q3)

Let him sleep (Go to Q4)

Check his BP (Go to Q2)

6. Stuart remains hypoglycaemic. This needs to be corrected first before he can have his breakfast.

What is the best course of action now?

Leave him to sleep (Go to Q4)

Check his BG (Go to Q3)

7. You run to fetch the Hypobox.

What is the best course of action now?

Give IV Dextrose (Go to Q8)

Give Glucotabs (Go to Q9)

Give Glucogel (Go to Q10)

8. Stuart is conscious so giving IV Dextrose is not the best course of action. Think again.

What is the best course of action now?

Give Glucotabs (Go to Q9)

Give Glucogel (Go to Q10)

9. You give 4-5 Glucotabs (15-20g). Then you retake Stuart's BG 10 mins later. Stuart's blood glucose levels are now 3.4 mmol/L

What is the best course of action now?

Give Glucogel (Go to Q10)

Give more Glucotabs (Go to Q12)

10. Giving Stuart Glucogel is not the best course of action. He can tolerate a less invasive treatment.

What is the best course of action now?

Give more Glucotabs (Go to Q12)

Give him his breakfast it is going cold (Q11)

11. Stuart is still hypoglycaemic; this must be corrected before he can have his breakfast.

What is the best course of action now?

Give more Glucotabs (Go to Q12)

12. You give Stuart a further 4-5 Glucotabs then retake his BG using the BG monitor. Stuart's blood glucose is now 4.5 mmol/L and he is feeling much better and now eats his breakfast. The doctor reviews his medication during his next round and discharges Stuart with Metformin and Gliclazide.

Is Stuart still at risk of Hypoglycaemia after discharge?

No (Go to Q14)

Yes (Go toQ13)

13. Stuart is still at risk of Hypoglycaemia because Gliclazide is a Sulphonyl urea (causes hypoglycaemia as a side effect).

Is Stuart still at risk of Hypoglycaemia after discharge?

Yes (Go to Q14)

14. Stuart is still at risk of Hypoglycaemia because Gliclazide is a Sulphonyl urea (causes hypoglycaemia as a side effect).

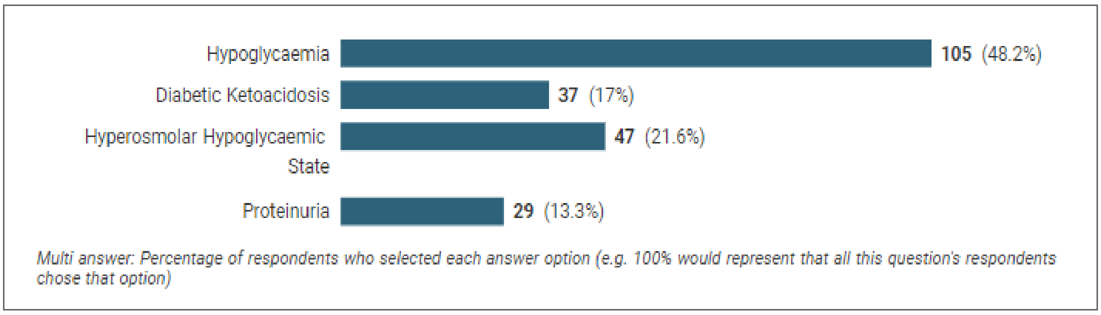
Is Stuart still at risk of Hypoglycaemia after discharge?

Yes (Go to Q15)

15. Stuart is still as risk of Hypoglycaemia after discharge because Gliclazide is a Sulphonylurea (causes hypoglycaemia as a side effect).

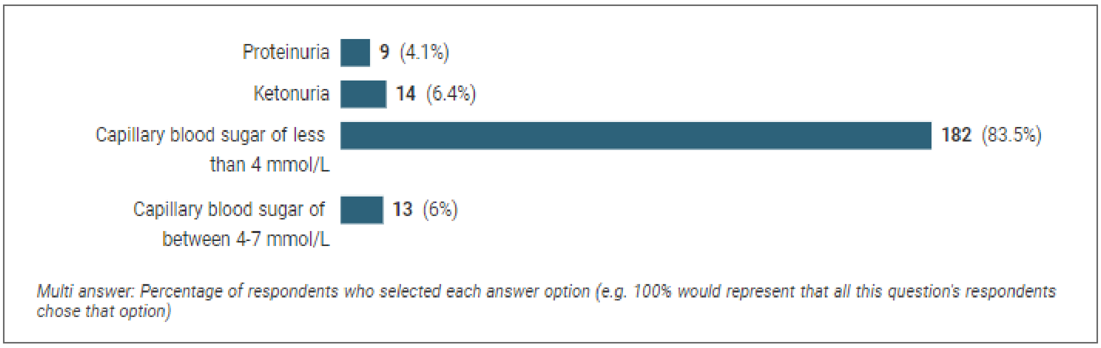
Appendix Three: Section of the Pre-Test

2 Very tight control of diabetes could lead to:



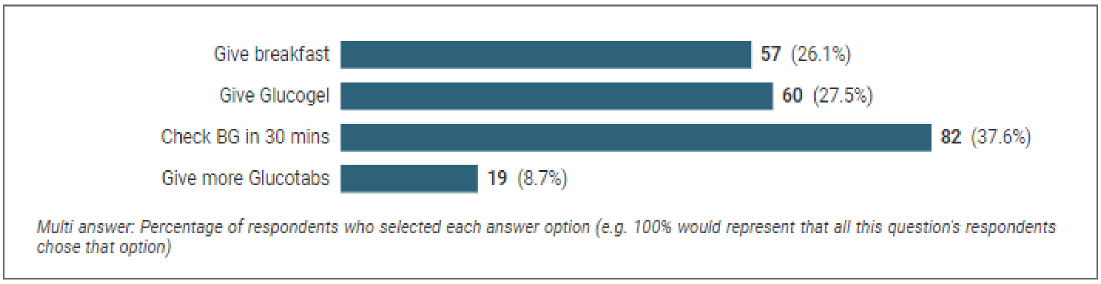
Multi answer: Percentage of respondents who selected each answer option (e.g. 100% would represent that all this question's respondents chose that option)

3 A diagnosis of hypoglycaemia is confirmed by:



Multi answer: Percentage of respondents who selected each answer option (e.g. 100% would represent that all this question's respondents chose that option)

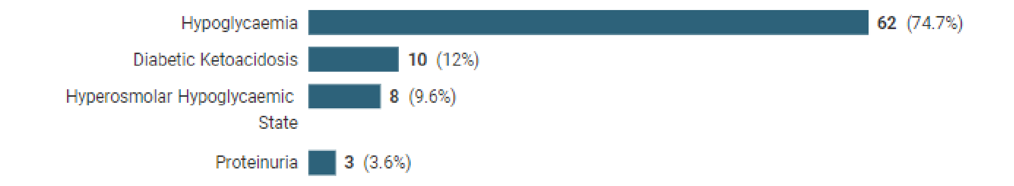
4 A patient has been suffering from hypoglycaemia and has already been given 4-5 Glucotabs. Their BG is now 3.4 mmol/L. What do you do now



Multi answer: Percentage of respondents who selected each answer option (e.g. 100% would represent that all this question's respondents chose that option)

Appendix Four: Section of the Experimental Group Post-Test

1 Very tight control of diabetes could lead to:

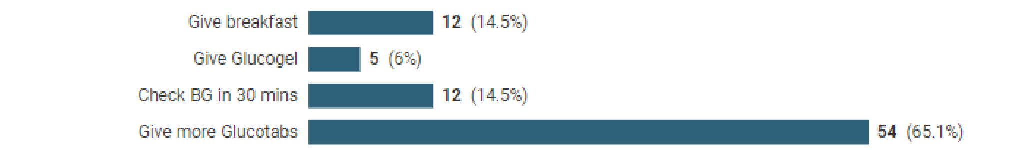


Multi answer: Percentage of respondents who selected each answer option (e.g. 100% would represent that all this question's respondents chose that option)

2 A diagnosis of hypoglycaemia is confirmed by:

Multi answer: Percentage of respondents who selected each answer option (e.g. 100% would represent that all this question's respondents chose that option)

3 A patient has been suffering from hypoglycaemia and has already been given 4-5 Glucotabs. Their BG is now 3.4 mmol/L. What do you do now?



Multi answer: Percentage of respondents who selected each answer option (e.g. 100% would represent that all this question's respondents chose that option)

Appendix Five: Focus Group Themes

Focus Group Results

Question 1. What was your experience of using the laptop case study?

Question 2 What were the challenges you faced?

Question 3. How do you think it could be used in the future?

Participant	Engagement		Confidence	Knowledge	Immersion	Challenges	Future
	Enjoyment	Ease of Use		Knowledge	Bringing it to life	Technical Issues	Future Uses
AC1	Obviously students will like to use this kind of technology for their learning.	I expect they would pick it up pretty quick.	If their theory practice gap has been closed they will begin to feel more confident.	I would hope that when they transfer that learning to a real situation, they would find it easier to translate that knowledge, that theory to the practice. So it is better for the theory-practice gap.	It combines the advantages of a simulated ward with those of a drama role play, in that you can have the sense of urgency as well as deteriorating vital signs in the patient. If the patient is getting anxious and you need to take their blood pressure you could make your … (avatar) talk to the patient to calm them down. It really does make you really focused on the situation and it makes it feel a lot more real	I think it is about investment, so if you are going to do this work in a HE setting you have got to have investment, not just in terms of money but also in time and appreciating the work people do.	There lots of uses, if we can straighten out the investment considerations. Hopefully it can be used lots more to help future students.

Appendix Six: CPMAO Configurations tested in this study

CPMAO configurations	Contextual Mechanisms	Programme Mechanisms	Agents	Outcome	Evaluation of the Configuration
CPMAO1#	**Amount of diabetic nursing experience and relevant computing EXPERIENCE required to complete exercise and understand complex diabetic concepts.**	Affordances of VR including opportunities to practise via experiential learning, and affordances of computing exercise, for example, immediate feedback given and the chance to repeat exercise	Confidence boosting via feedback	Improved knowledge of complex diabetic issues	This configuration **was not confirmed** through the qualitative and quantitative findings of this research.
CPMAO2#	Participant engagement level with the learning content	Affordances of VR, for example, visualisation of abstract concepts	Participant immersion with the desktop VR	Enjoyment of the learning activity	This configuration **was confirmed** through the qualitative and quantitative findings of this research.

Appendix Seven: Field notes

Systematic Review

2008 chosen as Gartner start of Rift

CASP quality chosen but not favoured by 1st reviewer

Own bespoke evaluation system queried but 2nd reviewer thought it was interesting as did editor.

I was told that my writing needs to be more interpretive and less descriptive.

Hypo quiz

Written with nurse specialist and Hampshire NHS trust. Glucogel question differed from trust to trust. Specifically, the words “very aggressive” were discussed and no consensus was reached. Dr James and Simone Penfold said go ahead with Glucogel answer. Questions were deliberately made to be challenging and were closely linked to the learning objectives of the students’ curriculum and the content of the diabetes case study (both the paper and VR versions).

Writing the case studies

Written with JJ and SP specialist nurses. We used Team-Based Leaning approach. For example, challenging and slightly ambiguous case, stressing justification for why students made the decisions they made. Originally drawn out on 3 pieces of sugar paper. Then turned into Visio decision tree document. The nurses and I checked the flow of the decision tree numerous times to eliminate errors.

Working with the software developer

Skype meetings times two to discuss and explain Visio decision tree. The LTS and I, then myself and the nurses checked the app for errors and improvements. Then several iterations of the app to and from the developer, myself and specialist nurses. Things that changed: look of senior nurse, BMI of patient (he needed to look heavier), skin colour of patient, blanket on patient’s legs. We also moved the location of the Hypobox as we thought it should be near the senior nurse. We added a countdown to the Glucose monitor. I asked for analytics to be added so that the lecturer could see how well their students were doing.

Finally, a piloting team comprising of CEL colleagues (including two with a medical background), two nurses and a software developer piloted the app. I asked the developer to produce a short video of the app so I could use it at conferences.

There were several modifications we would like the app to have in the future, including: real time graphing capabilities to show which choices students are making. Animations for the monitoring equipment. For the patient to sit up and eat at the end. For the senior nurse to be in the corridor. For the nurse avatar to leave the room to fetch the hypo box. For teamwork option that allows the online chat option for communication.

Writing the Survey

TAM and Immersive tendencies were used to base questions on to aid reliability. A pilot could not take place as the students were on their summer holidays. They were asked but the response rate was 0. However, the statistician felt that tried and tested questions are more reliable than piloting a survey.

I used the option of “an answer is required” this meant that all the respondents’ surveys could be used. All surveys were complete and could move on to analysis phase. At the ethics stage I asked if I could ask for the student’s university number. This meant that I could trace the students through and see how their answers to the questions were related to their performance on the MCQ, and so forth,

Implementing the pre-survey

Contacted the students by group email to introduce myself and the project and ask them to bring a computing device to their next lecture in order that they could complete the online consent and survey if they chose to. I printed out paper versions in case of any Internet problem. However, all participants chose to complete the survey online. I went to see the students in person, to introduce the project and assist them with logging into the online survey.

Several academics encouraged participation in that they introduced me and stressed the importance of such research. Unfortunately, one lecture over ran. This meant that the participants had to complete the survey at home (due to this I lost some potential participants). I realised that when I held online survey completion sessions at the end of the day, more students were inclined to leave (especially the back row), and again this led to a reduction in participant numbers.

I explained that everyone would have access to the VR training programme after the data collection phase, if they requested access.

Mental Health Pre-survey

The lecturer stayed to the end. Uptake was good.

The VR Diabetes Case study Session

Randomisation

Students were randomized into two groups (VR desktop, paper-based case study). They were then contacted email and provided with login details and a software download link. They were told what group they would be in on the study day so that they could bring the correct device to run the software. If they did not bring a laptop, they went in Control Group but stated on the survey why they had not brought the laptop, and so forth, no students completed both the laptop and paper versions. One student who was supposed to be in the laptop group arrived late, but as they had seen the hypoglycaemia input, they could not complete the computer-based case study.

BU session

Students who completed the paper-based version were asked to complete it twice so that they could try to improve their score. The VR students were given the same instructions. Students who completed the computer version had a good play and were enjoying themselves. Not sure how well they concentrated on the nursing content? This was since they had to share one machine between three, one after the other. The Control Group were quieter and concentrated better.

Students who had not consented to take part in the study completed the paper-based case study but did not go on to complete the online survey and post quiz.

Several students were stuck in traffic. Lots of students had been absent from the induction day. Several were not on group email system but joined Control Group on the day (completed the pre-survey). Only two had followed instruction to bring laptop with software on. A few others had laptops but had not downloaded the software. I immediately sent another reminder to upcoming groups to bring laptops and download software. I asked CEL team if we had any spare laptops (I had used my own and two from CEL in the morning). I acquired a third CEL laptop for the afternoon. I then downloaded the software for the afternoon. I also photocopied more paper case studies. The ones who completed the simulation seemed to enjoy it. The paper case study people not in study reported it was difficult to complete. The Control Group were fine. All groups were asked to complete several times. Many students had failed the unit, and this impacted on numbers and morale of the students.

Yeovil Session

I had reminded them to bring laptops. Lots had laptops and were enthusiastic, but the download was very slow. Some had downloaded in advance (as per my instructions to n=70). Several were able to complete using the laptops we took with us. Overall the room had a positive atmosphere and the students seemed to enjoy the computer version. One lady had completed it at home and had really enjoyed it. On this occasion we were able to solve the problem of too many avatars in the room (by switching off the Wi-Fi).

Overall, absence was a problem for the pre and post-tests. Many students completed the paper or computer version plus post survey at home. They had not had diabetic input by that time. Judging by their scores no one took the opportunity to consult google or a textbook. In fact, the chance of asking a friend was lowered when working from home.

Mental Health session

3 LTs present. Lecturer not present. Each student had paper instructions for laptop or the paper case. Link to end survey was on the same piece of paper. Everyone completed the post survey with no confusion. Instructions for what to do when finished so everyone happy.

Implementing the Focus Groups

I asked students who had completed the survey if they would participate in a focus group session. I carefully booked the focus group sessions in very close to the students' lectures so that they would already be on campus. I made room bookings not far from the lecture room bookings. Student's received a meal voucher for taking part. Chocolate was put on tables.

4th October 13:00

A group of 5 students attended. One was a mature student. One was Portuguese. It was a very balanced discussion starting with the technological difficulties of having more than one avatar. One student discussed how the paper version was quick and easy to use and discussed usability issues of the computer version. Then then discussion turned to the advantages of the computer version. Students discussed how it would help as a revision aid and thought it would be good to have more scenarios that included video and more of a teaching element as well. The atmosphere was supportive and positive throughout.

5th October. 11:15

One student came by herself. She felt that using it in the lecture theatre first would be of benefit if you are not techy. She found quite a few glitches in the software but enjoyed it and felt it would help her learning. She was a mature student.

6th October. 12:15

Two students attended. Both mature students. One was also a HCA. Both really enjoyed it and made positive comments. They found it easy to download and use. They described how when they were in the simulated ward they were looking for monitors to be where they were on the ward.

29th Feb 2019

X6 Mental health Students stayed behind after the study. Very positive. Learning to move and downloading only drawbacks. Found very useful. Text boxes useful for dyslexic and visual learners. Had hoped you could pick up items from Hypobox. Mac- pink hair. Would like more examples for high risk mental health emergencies.

Appendix Eight: Ethics Protocol

Bournemouth University

Research Ethics Checklist

About Your Checklist	
Reference Id	21833
Status	Approved
Date Approved	23/07/2018 10:36:13
Date Submitted	25/06/2018 12:58:26

Researcher Details	
Name	Heidi Singleton
Faculty	Faculty of Media & Communication
Status	Postgraduate Research (MRes, MPhil, PhD, DProf, EngD, EdD)
Course	Postgraduate Research - FMC
Is This External Funding?	No
Please list any persons or institutions that you will be conducting joint research with, both internal to BU as well as external collaborators.	RBCH

Project Details	
Title	Evaluation of a virtual reality diabetes education innovation for healthcare students
End Date of Project	31/08/2020
Proposed Start Date of Data Collection	02/09/2018
Original Supervisor	Deborah Holley
Approver	Research Ethics Panel

Summary - no more than 500 words (including detail on background methodology, sample, outcomes, etc.)

The research has a strong rationale in view of the rapid evolution of VR technologies, which has not been accompanied by the appraisal of their effectiveness for student-centred learning. Immersive virtual worlds such as SL have been deployed in HE for the last decade, with research indicating strong student engagement (Philips et al., 2015), but often relying on self-report for learning outcomes (Kurilovas et al., 2016). This project will bridge the gaps of effective evaluation of VR by implementing original, empirical research exploring theoretical instructional design and evaluation of these technologies for the student body; with a plan for scalability across contexts.The teaching and learning exercise will be deployed operationalising the research gaps found in the literature review. It is expected that the exercise will utilise virtual reality technology with googlecardboard headsets. The participant group will be BU Adult and Mental Health Nursing students. Process and outcome evaluations of the intervention exercise design will include qualitative measures including:•video-analysis of implementation sessions (non-verbal, implicit level)•questionnaires for student engagement (approx 200 students at BU and 50 students at the Yeovil site) and questionnaire to staff to establish barriers and enablers to using VR I their teaching (approx 50 participants)•phenomenographic interviews with approximately 15-20 students and academics (focusing on the limited possible ways of experiencing a given phenomenon, across a group of individuals) •focus groups sessions involving participants meeting in a virtual world to discuss their experience of the intervention exercise. -quantitative measures including:-attendance/submission/assessment records pre and post teaching and learning exercise•responses to before/after surveys. The intervention results, in turn, should influence pedagogical practices and methods in education in general, and a dissemination pathway with practitioners will be designed to this effect. It is expected that the model generated by this project can be extended to improve the end user experience of virtual, augmented and

mixed reality users beyond the lecture theatre, with transferability to different learning contexts for example: nursing homes, hospitals and the community.

External Ethics Review	
Does your research require external review through the NHS National Research Ethics Service (NRES) or through another external Ethics Committee?	No

Research Literature	
Is your research solely literature based?	No

Human Participants	
Does your research specifically involve participants who are considered vulnerable (i.e. children, those with cognitive impairment, those in unequal relationships—such as your own students, prison inmates, etc.)?	No
Does the study involve participants age 16 or over who are unable to give informed consent (i.e. people with learning disabilities)? NOTE: All research that falls under the auspices of the Mental Capacity Act 2005 must be reviewed by NHS NRES.	No
Will the study require the co-operation of a gatekeeper for initial access to the groups or individuals to be recruited? (i.e. students at school, members of self-help group, residents of Nursing home?)	No
Will it be necessary for participants to take part in your study without their knowledge and consent at the time (i.e. covert observation of people in non-public places)?	No
Will the study involve discussion of sensitive topics (i.e. sexual activity, drug use, criminal activity)?	No
Are drugs, placebos or other substances (i.e. food substances, vitamins) to be administered to the study participants or will the study involve invasive, intrusive or potentially harmful procedures of any kind?	No
Will tissue samples (including blood) be obtained from participants? Note: If the answer to this question is 'yes' you will need to be aware of obligations under the Human Tissue Act 2004.	No
Could your research induce psychological stress or anxiety, cause harm or have negative consequences for the participant or researcher (beyond the risks encountered in normal life)?	No
Will your research involve prolonged or repetitive testing?	No
Will the research involve the collection of audio materials?	No
Will your research involve the collection of photographic or video materials?	Yes
Will financial or other inducements (other than reasonable expenses and compensation for time) be offered to participants?	No
Please explain below why your research project involves the above mentioned criteria (be sure to explain why the sensitive criterion is essential to your project's success). Give a summary of the ethical issues and any action that will be taken to address these. Explain how you will obtain informed consent (and from whom) and how you will inform the participant(s) about the research project (i.e. participant information sheet). A sample consent form and participant information sheet can be found on the Research Ethics website.	
The virtual realty implementation sessions will be recorded using the usual lecture capture (Panopto). Student nurses who will be in the lectures will be given a participant information sheet and a consent form. They will be informed that they do not have to take part in the study and can opt out at any time without having to provide a reason.	

Approved Amendments

Message	Hello, I have collected data for the laptop arm of my RCT. My app is now ready.Please can I collect data for the app version with 8- 2nd year nurses in Sept/october 2019. All of the surveys etc will be the same it is just another group of students to extend the study.
Date Submitted	22/03/2019 10:18
Comment	

Date Approved	25/03/2019 11:35
Approved By	Suzy Wignall

Final Review

Will you have access to personal data that allows you to identify individuals OR access to confidential corporate or company data (that is not covered by confidentiality terms within an agreement or by a separate confidentiality agreement)?	No
Will your research involve experimentation on any of the following: animals, animal tissue, genetically modified organisms?	No
Will your research take place outside the UK (including any and all stages of research: collection, storage, analysis, etc.)?	No

Please use the below text box to highlight any other ethical concerns or risks that may arise during your research that have not been covered in this form.

Participants will be informed of the very slight risk of nausea that can be caused by virtual reality use. Headsets will be brand new to start with and will then be cleared with sterilising wipes before the next use. Participants will be sat down whilst using VR to prevent them from tripping up. Facilitators will be on hand to help.

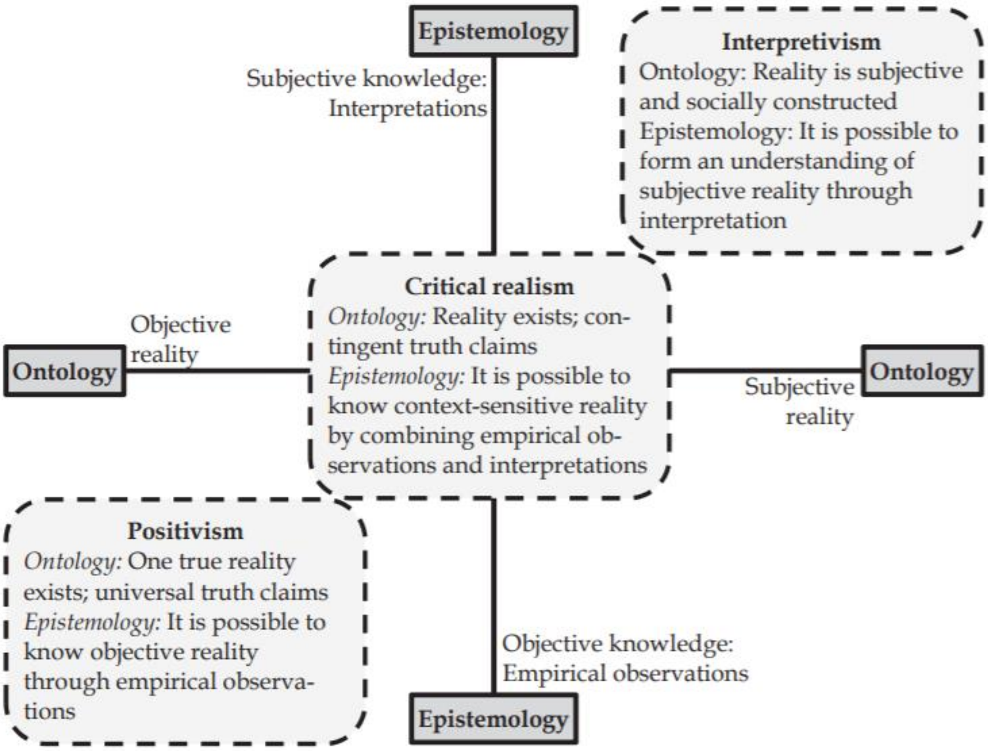
Attached documents

updated participant agreement for oct 2018.docx - attached on 10/10/2018 14:45:25

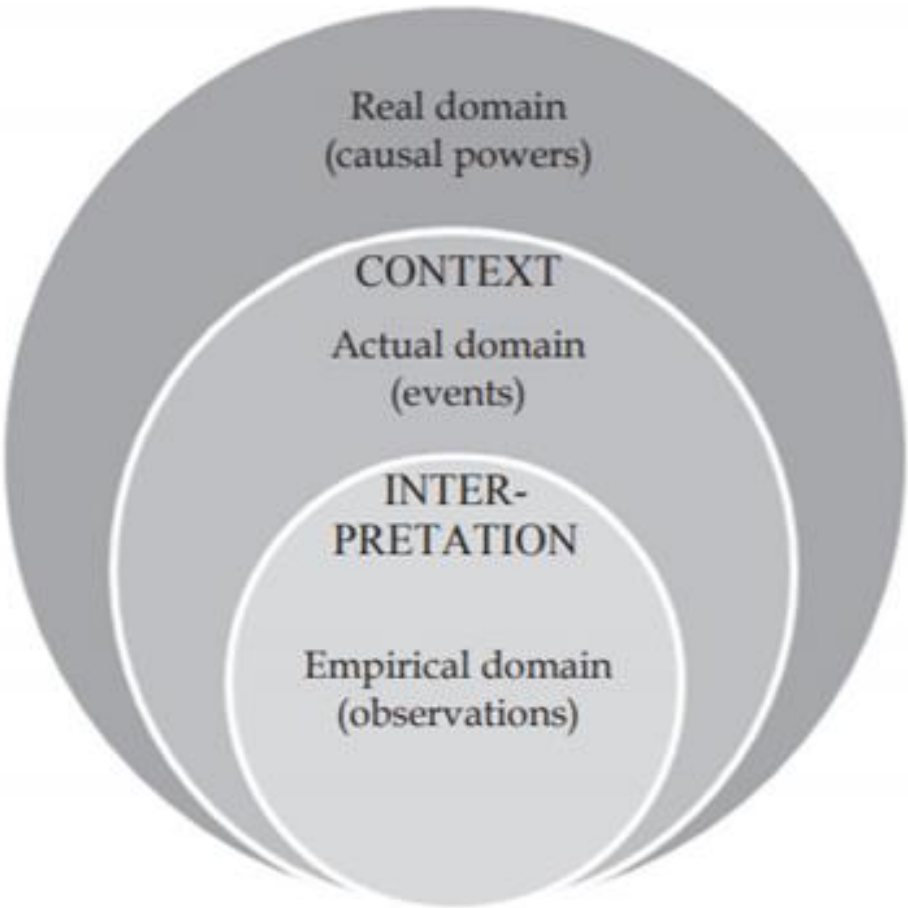
Approved Amendments

Message	Please can the ethical clearance be extended to include 2nd Year Mental Health Nurses as well as 2nd Year Adult Nurses (already had clearance for this). The Mental Health Nurses receive the same anatomy and physiology lectures in their 1st Year as the Adult Nurses. They also have the same diabetes input from Dr Janet James in their second year. Therefore, it seems sensible to include them in the Virtual Reality Diabetes study and offer them the same opportunities that the Adult Nurses receive.
Date Submitted	10/10/2018 14:13
Comment	
Date Approved	12/10/2018 09:30
Approved By	Suzy Wignall

Appendix Nine: Ontological and epistemological comparisons of research paradigms Adapted from (Järvensivu and Törnroos, 2010).



Appendix Ten: Ontological layers of CR- based upon (Bhaskar, 1989)



References

ACHUTHAN, K., SANEESH, F. & SHYAM, D. 2017. Augmented reflective learning and knowledge retention perceived among students in classrooms involving virtual laboratories. *Education & Information Technologies,* 22**,** 2825-2855.

ADER, H., MELLENBERG, G. & HAND, D. 2008. Advising on Research Methods: A Consultant's companion. Germany: Johannes van Kessel Publishing.

AHMED, B. & KHAN, M. Hypoglycemia: Its effect on patients with diabetes 2019. Omar Bin Khattab Health Center.

AKCAYIR, M. & AKCAYIR, G. 2017. Advantages and challenges associated with AR for education: a systematic review of the literature. *Educational research review,* 20**,** 1-11.

ALDERSON, P. 2020. Powerful knowledge and the curriculum: Contradictions and dichotomies. *British Educational Research Journal,* 46**,** 26-43.

ALHABASH, S., ALMUTAIRI, N., LOU, C. & KIM, W. Pathways to Virality: Psychophysiological Responses Preceding Likes, Shares, Comments, and Status Updates on Facebook. 2019 2019 Great Britain. Taylor & Francis, 196.

ALI, K., KISIELEWSKA, K., SUBHAN, M. & TREDWIN, C. 2020. How does dyslexia impact on the educational experiences of healthcare students? A qualitative study. *European Journal of Dental Education,* 24**,** 154.

ALLANA, S. & CLARK, A. 2018. Applying Meta-Theory to Qualitative and Mixed-Methods Research: A Discussion of Critical Realism and Heart Failure Disease Management Interventions Research. *International Journal of Qualitative Methods,* 17**,** 1-9.

ALLCOAT, D. & VON MÜHLENEN, A. 2018. Learning in virtual reality: Effects on performance, emotion and engagement. *Research in Learning Technology***,** 1.

AMIN, A., GROMALA, D., TONG, X. & SHAW, C. 2016. Immersion in Cardboard VR Compared to a Traditional Head-Mounted Display. International Conference on Virtual, Augmented and Mixed Reality VAMR 2016 : Virtual, Augmented and Mixed Reality

ARBAUGH, J. B. & DURAY, R. 2002. Technological and Structural Characteristics, Student Learning and Satisfaction with Web-based

Courses: An Exploratory Study of Two On-line MBA Programs. Great Britain: SAGE PUBLICATIONS.

ARCHER, M., LAWSON, T., NORRIE, A., BHASKAR, R. & COLLIER, A. 2013. *Critical Realism : Essential Readings,* Hoboken, Routledge.

ARCHER, M. S. 2010. CRITICAL REALISM AND RELATIONAL SOCIOLOGY Complementarity and Synergy. *Journal of Critical Realism,* 9**,** 199.

ARVANITIS, T., PETROU, A., KNIGHT, J., SAVAS, S., SOTIRIOU, S., GARGALAKOS, M. & GIALOURI, E. 2009. Human factors and qualitative pedagogical evaluation of a mobile augmented reality system for science education used by learners with physical disabilities. *Personal & Ubiquitous Computing,* 13**,** 243-250.

ASPINALL, C., JACOBS, S. & FREY, R. 2019. Intersectionality and Critical Realism: A Philosophical Framework for Advancing Nursing Leadership. *Advances in Nursing Science,* 42**,** 289-296.

ASTBURY, B. & LEEUW, F. L. 2010. Unpacking Black Boxes: Mechanisms and Theory Building in Evaluation. *American Journal of Evaluation,* 31**,** 363-381.

ATA, R. 2016. An Exploration of Higher Education Teaching in Second Life in the Context of Blended Learning. *Turkish Online Journal of Educational Technology - TOJET,* 15**,** 9-26.

ATHANASAKIS, E. 2019. A meta-synthesis of how registered nurses make sense of their lived experiences of medication errors. *Journal of Clinical Nursing,* 28**,** 3077-3095.

ATTARAN, M. & MORFN-MANIBO, R. 2018. Your future reality will be digital. *ISE: Industrial & Systems Engineering at Work,* 50**,** 26-31.

AVKIRAN, N. K. 2018. An in-depth discussion and illustration of partial least squares structural equation modeling in health care. *Health Care Management Science,* 21**,** 401-408.

AZUMA, R. 1997. A survey of augmented reality. *Presence: Teleoperators & Virtual Environments,* 6**,** 355.

BACCA, J., BALDIRIS, S., FABREGAT, R., GRAF, S. & KINSHUK 2014. Augmented Reality Trends in Education: A Systematic Review of Research and Applications. *Journal of Educational Technology & Society***,** 133.

BAILEY, J. H. & WITMER, B. G. Learning and Transfer of Spatial Knowledge in a Virtual Environment. 1994. The Society, 1158.

BAKER, M. A., ROBINSON, J. S. & KOLB, D. A. 2012. Aligning Kolb's Experiential Learning Theory with a Comprehensive Agricultural Education Model. *Journal of Agricultural Education,* 53**,** 1-16.

BAKIR, N. 2016. Technology and Teacher Education: A Brief Glimpse of the Research and Practice That Have Shaped the Field. *TechTrends: Linking Research and Practice to Improve Learning,* 60**,** 21-29.

BARNETT, S. G., GALLIMORE, C. E., PITTERLE, M. & MORRILL, J. 2016. Impact of a Paper vs Virtual Simulated Patient Case on Student-Perceived Confidence and Engagement. *American Journal of Pharmaceutical Education,* 80**,** 1-10.

BARRETT, J. 2004. Side effects of virtual environments: A review of the literature. Edinburgh, Australia: Defense sciences and technology organisation information sciences lab.

BAXTER, G. & HAINEY, T. 2019. Student perceptions of virtual reality use in higher education. *Journal of Applied Research in Higher Education*.

BAYRAM, S. B. & CALISKAN, N. 2019. Effect of a game-based virtual reality phone application on tracheostomy care education for nursing students: A randomized controlled trial. *Nurse Education Today,* 79**,** 25-31.

BAZZAZA, M., ALZUBAIDI, M. & ZEMERLY, M. 2016. Impact of smart immersive mobile learning in language literacy education. IEEE.

BEHMKE, D., BRANNOCK, E. & KERVEN, K. AR Chemistry: An Undergraduate, Technology-Based Research and Development Initiative To Incorporate AR Molecular Models in the Chemistry Curriculum. Technology Integration in Chemistry Education and Research (TICER), 2019.

BELL, M. 2008. Toward a definition of VWs. *Journal of VW Research,* 1**,** 7-29.

BELLASS, S., BALMER, A., MAY, V., KEADY, J., BUSE, C., CAPSTICK, A., BURKE, L., BARTLETT, R. & HODGSON, J. 2019. Broadening the debate on creativity and dementia: A critical approach. *Dementia: The International Journal of Social Research and Practice,* 18**,** 2799-2820.

BENTON, T. & CRAIB, I. 2005. Philosophy of Social Science. The Philosophical Foundations of Social Thought. *Revista Española de Investigaciones Sociológicas (REIS)***,** 270.

BERG, L. & VANCE, J. 2017. Industry use of virtual reality in product design and manufacturing: a survey. *Virtual Reality,* 1**,** 1-17.

BHASKAR, R. 1989. *The possibility of naturalism: a philosphical critique of the contemporary human sciences.,* Harvester Wheatsheaf, London.

BIRT, J., STROMBERGA, Z., COWLING, Z. & MORO, C. 2018. Mobile Mixed Reality for Experiential Learning and Simulation in Medical and Health Sciences Education. *Information* 9.

BLACK, P. & WILIAM, D. 1998. Inside the Black Box: Raising Standards through Classroom Assessment. *The Phi Delta Kappan,* 80**,** 139.

BLACKWOOD, B., O'HALLORAN, P. & PORTER, S. 2010. On the problems of mixing RCTs with qualitative research. *Journal of Research in Nursing,* 15**,** 511-521.

BLAIKIE, N. W. H. & PRIEST, J. 2017. *Social research. [electronic resource] : paradigms in action*, Polity.

BLOM, B. & MORÉN, S. 2011. ANALYSIS OF GENERATIVE MECHANISMS. *Journal of Critical Realism,* 10**,** 60.

BOET, S., BOULD, M. D., SHARMA, B., REVEES, S., NAIK, V. N., TRIBY, E. & GRANTCHAROV, T. 2013. Within-Team Debriefing Versus Instructor-Led Debriefing for Simulation-Based Education: A Randomized Controlled Trial. United States: Lippincott Williams and Wilkins.

BOLLEN, K. A. & BAULDRY, S. 2011. Three Cs in Measurement Models: Causal Indicators, Composite Indicators, and Covariates. United States: AMERICAN PSYCHOLOGICAL ASSOCIATION.

BONELL, C., WARREN, E., FLETCHER, A. & VINER, R. 2016. Realist trials and the testing of context-mechanism-outcome configurations: a response to Van Belle et al. *Trials,* 17**,** 478-478.

BOOTH, A. 2016. Qualitative Evidence Synthesis. *Patient Involvement in Health Technology Assessment***,** 187-199.

BOWER, M., LEE, M. & DALGARNO, B. 2017. Collaborative learning across physical and virtual worlds: Factors supporting and constraining learners in a blended reality environment. *British Journal of Educational Technology,* 48**,** 407-430.

BRANNEY, J. & PRIEGO-HERNÁNDEZ, J. 2018. A mixed methods evaluation of team-based learning for applied pathophysiology in undergraduate nursing education. *Today, Nurse Education,* 61**,** 127–133.

BRANT, T. 2001. Consequences of Realism for Sociological Theory-Building. *Journal for the Theory of Social Behaviour***,** 167.

BRIDGE, P., GUNN, T., KASTANIS, L., PACK, D., ROWNTREE, P., STARKEY, D., MAHONEY, G., BERRY, C., BRAITHWAITE, V. & WILSON-STEWART, K. 2014. The development and evaluation of a medical imaging training immersive environment. *Journal of Medical Radiation Sciences,* 61**,** 159-165.

BRIGHAM, T., J. 2017. Reality Check: Basics of Augmented, Virtual, and Mixed Reality. *Medical Reference Services Quarterly* 36**,** 171-178.

BROWN, E., HOBBS, M. & GORDON, M. 2008. A virtual world environment for group work. *International Journal of Web-Based Learning and Teaching Technologies,* 3**,** 1-12.

BRYMAN, A. 2009. Mixed methods in organizational research. *In:* BUCHANAN, D. A. & BRYMAN, A. (eds.) *The Sage handbook of organizational research methods.* Thousand Oaks, CA: Sage Publications Ltd.

BUTT, A. L., KARDONG-EDGREN, S. & ELLERTSON, A. 2018. Using Game-Based Virtual Reality with Haptics for Skill Acquisition. *Clinical Simulation in Nursing,* 16**,** 25-32.

BYRNE, D. & UPRICHARD, E. Useful Complex Causality. 2012. Oxford; Oxford University Press, 109.

CAMPBELL, K., MASSEY, D., BROADBENT, M. & CLARKE, K. A. 2019. Factors influencing clinical decision making used by mental health nurses to provide provisional diagnosis: A scoping review. *International Journal of Mental Health Nursing,* 28**,** 407-424.

CANTOR, G. N. 1982. The Possibility of Naturalism: A Philosophical Critique of the Contemporary Human Sciences. Roy Bhaskar. The University of St. Andrews.

CARBO, A. R. & HUANG, G. C. Promoting clinical autonomy in medical learners. 2019 2019 Great Britain. John Wiley & Sons Ltd, 454.

CARMIGNIANI, J., FURHT, B., ANISETTI, M., CERAVOLO, P., DAMIANI, E. & IVKOVIC, M. 2011. Augmented reality technologies, systems and applications. *Multimedia Tools & Applications,* 51**,** 341.

CASP 2013. Critical Appraisal Skills Programme : Qualitative research checklist.

CASTILLO, B., IVÁN, R., SÁNCHEZ, C., GUADALUPE, V., VILLEGAS, V. & OSIRIS, O. 2015. A Pilot Study on the Use of Mobile Augmented Reality for Interactive Experimentation in Quadratic Equations. *Mathematical Problems in Engineering,* 2015**,** 1-13.

CHAN, M. & ZANG, Y.-L. 2007. Nurses' perceived and actual level of diabetes mellitus knowledge: results of a cluster analysis. *Journal of CLinical Nursing,* 16.

CHANG, X., ZHANG, D. & DIN, X. 2016. Application of Virtual Reality Technology in Distance Learning. *International Journal of Emerging Technologies in Learning,* 11**,** 76.

CHAU KWONG, W., RAFTERY, J. & WALKER, A. 1998. The baby and the bathwater: research methods in construction management. *Construction Management & Economics,* 16**,** 99-104.

CHEAH, U. 2017. The smartphone way to see a tropical jungle; ArtScience Museum presents first augmented and virtual reality exhibition in Asia.

CHEN, C. & TSAI, Y. 2012. Interactive augmented reality system for enhancing library instruction in elementary schools. *Computers & Education,* 59**,** 638-652.

CHEN, Y. L. 2016. The Effects of Virtual Reality Learning Environment on Student Cognitive and Linguistic Development. *Asia-Pacific Education Researcher (Springer Science & Business Media B.V.),* 25**,** 637-646.

CHENG, K. & TSAI, C. 2013. Affordances of Augmented Reality in Science Learning: Suggestions for Future Research. *Journal of Science Education & Technology,* 22**,** 449-462.

CHEONG, D. 2010. The effects of practice teaching sessions in second life on the change in pre-service teachers' teaching efficacy. *Computers and Education,* 55**,** 868-880.

CHILDS, M. 2010. A conceptual framework for mediated environments. *Educational Research,* 52**,** 197-213.

CHINIARA, G., CLARK, M., JAFFRELOT, M., POSNER, G. & RIVIERA, E. 2019. Education, Operations, and Engineering. *In:* CHINIARA, G. (ed.) *Clinical Simulation (Second Edition).* Canada: Elsevier.

CHODOS, D., STROULIA, E., KING, S. & CARBONARO, M. 2014. A framework for monitoring instructional environments in a virtual world. *British Journal of Educational Technology,* 45**,** 24-35.

CLAMAN, F. 2015. The impact of multiuser virtual environments on student engagement. *Nurse Education in Practice,* 15**,** 13-16.

CLARKE, V. & BRAUN, V. 2018. Using thematic analysis in counselling and psychotherapy research: A critical reflection. *Counselling & Psychotherapy Research,* 18**,** 107-110.

COBBETT, S. & SNELGROVE-CLARKE, E. 2016. Virtual versus face-to-face clinical simulation in relation to student knowledge, anxiety, and self-confidence in maternal-newborn nursing: A randomized controlled trial. *Nurse Education Today,* 45**,** 179-184.

COHEN, L., MANION, L. & MORRISON, K. 2017. *Research Methods in Education,* Routledge.

COLEMAN, P. 2019. An Examination of Positivist and Critical Realist Philosophical Approaches to Nursing Research. *International Journal of Caring Sciences,* 12**,** 1218-1224.

COLLIER, A., HODGINS, M., JOHNSTON, B. & TIEMAN, J. 2019. Evidence-based palliative care: How can we account for the messy world of practice? *Palliative Medicine,* 33**,** 723-725.

COLLINS, H., EVANS, R., WEINEL, M., LYTTLETON-SMITH, J., BARTLETT, A. & HALL, M. 2017. The imitation game and the nature of mixed methods. *Journal of Mixed Methods Research,* 11**,** 510-527.

CONSORTI, F., MANCUSO, R., NOCIONI, M. & PICCOLO, A. 2012. Efficacy of virtual patients in medical education: A meta-ansysis of randomised studies. *Computers and education,* 59**,** 1001-1008.

COOK, D. 2013. Mastery Learning for Health Professionals Using Technology-Enhanced Simulation

A Systematic Review and Meta-Analysis.

COOKE, A., SMITH, D. & BOOTH, A. 2012. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. *Qual Health Res,* 22**,** 1435-43.

CORRY, M., PORTER, S. & MCKENNA, H. 2019. The redundancy of positivism as a paradigm for nursing research. *Nursing Philosophy: An International Journal For Healthcare Professionals,* 20**,** e12230-e12230.

COSTA, V. & MONTEIRO, S. 2018. From Potential Absorptive Capacity to Knowledge Creation in Organisations: The Mediating Role of Knowledge Storage and Realised Absorptive Capacity. *Journal of Information & Knowledge Management,* 17**,** 1-16.

CRESWELL, J. W. 2018. *Research Design: Qualitative, quantitative and mixed methods approaches, Sage publications*.

CRUICKSHANK, J. 2004. A tale of two ontologies: an immanent critique of critical realism. Great Britain: Blackwell Publishing Ltd.

CRUZ-NEIRA, C., FERNANDEZ, M. & PORTALES, C. 2018. Virtual Reality Games. Open Access Multimodal Technologies Interact.

CSIKSZENTMIHALYI, M. The Contribution of Flow to Positive Psychology. 2000 2000. Philadelphia; Templeton Foundation Press; c, 387.

CUGELMAN, B. 2013. Gamification: what it is and why it matters to digital health behavior change developers. *JMIR Serious Games,* 1**,** e3-e3.

DALGARNO, B. & LEE, M. 2010. What are the learning affordances of 3-D virtual environments? *Part of a special issue: Crossing boundaries: Learning and teaching in virtual worlds,* 41**,** 10-32.

DALKIN, S. M., GREENHALGH, J., JONES, D., CUNNINGHAM, B. & LHUSSIER, M. 2015. What's in a mechanism? Development of a key concept in realist evaluation. *Implementation Science,* 10**,** 1-7.

DANERMARK, B. 2002. *Explaining Society : An Introduction to Critical Realism in the Social Sciences,* London, Routledge.

DANG, N., TAVANTI, T., RANKIN, I. & COOPER, M. 2009. A comparison of different input devices for a 3D environment. *INTERNATIONAL JOURNAL OF INDUSTRIAL ERGONOMICS***,** 554.

DASS, S., DABBAGH, N. & CLARK, K. 2011. Using VWs what the research says. *Quarterly Review of Distance Education,* 12**,** 95-111.

DAVIS-REYES, S., STILLSMOKING, K. & CHADWICK-HOPKINS, D. 2008. Implementation and Evaluation of a Virtual Simulator System: Teaching Intravenous Skills. *Clinical Simulation in Nursing,* 4**,** 43-49.

DAVIS, F., BAGOZZI, R. & WARSHAW, P. 1989. User Acceptance of Computer Technology. *Management Science,* 35.

DE LEÓN, L. 2013. The Clash of Epistemologies: A Study of the Transformative Learning of Bilingual Pre-service Teachers Engaged in Simulations in a Virtual Environment. *HETS Online Journal,* 4**,** 240-267.

DEPAPE, A., *BARNES, M.* & *PETRYSCHUK, J.* 2019. Students' Experiences in Higher Education With Virtual and Augmented Reality: A Qualitative Systematic Review. *Innovative Practice in HE.,* 3.

DIAMANTOPOULOS, A., SARSTEDT, M., FUCHS, C., WILCZYNSKI, P. & KAISER, S. 2012. Guidelines for choosing between multi-item and single-item scales for construct measurement: a predictive validity perspective. Great Britain: Springer Science + Business Media.

DOHERTY-RESTREPO, J. L., HARRELSON, K. E., SWINNIE, T. & MONTALVO, A. M. 2017. Does simulation-based training increase athletic training students' clinical confidence and competence in performing a cardiovascular screening? *Journal of Allied Health,* 46**,** 171-178.

DOLPHIN, S. & O'CONNOR, E. A Graduate School Residency in a Virtual Reality Setting. World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education, 2019 USA. Association for the Advancement of Computing in Education (AACE).

DOOLEY, C. M., CALANDRA, B. & HARMON, S. 2014. Preservice Teachers Experience Reading Response Pedagogy in a Multi-User Virtual Environment. *Journal of Digital Learning in Teacher Education,* 30**,** 121-130.

DOWNWARD, P. & MEARMAN, A. 2007. Retroduction as mixed-methods triangulation in economic research: reorienting economics into social science. Great Britain: Oxford University Press.

DREHER, C., REINERS, T., DREHER, H. & DREHER, N. 2009. Virtual worlds as a context suited for information systems education: discussion of pedagogical experience and curriculum design with reference to second life. *Journal o Information Systems Education.,* 20**,** 211-224.

DRLJEVIĆ, N., WONG, L. & BOTIČKI, I. 2017. Where Does My Augmented Reality Learning Experience (ARLE) Belong? A Student and Teacher Perspective to Positioning ARLEs. *IEEE Transactions on Learning Technologies, Learning Technologies, IEEE Transactions on, IEEE Trans. Learning Technol.***,** 419.

DUBOVI, I., LEVY, S. & DAGAN, E. 2017. Now I know how! The learning process of medication administration among nursing students with non-immersive desktop virtual reality simulation. *Computers and Education,* 113**,** 16-27.

DUNCAN, I., MILLER, A. & JIANG, S. 2012. A taxonomy of virtual worlds usage in education. *British Journal of Educational Technology,* 43**,** 949-964.

DUNLEAVY, M., DEDE, C. & MITCHELL, B. 2009. Affordances and Limitations of Immersive Participatory Augmented Reality Simulations for Teaching and Learning. *Journal of Science Education and Technology***,** 7.

DYER, E., SWARTZLANDER, B. J. & GUGLIUCCI, M. R. 2018. Using virtual reality in medical education to teach empathy. *Journal of the Medical Library Association,* 106**,** 498-500.

EACOTT, S. 2015. Studying Administration Relationally. *Educational Leadership Relationally***,** 63.

EASTON, G. 2010. Critical realism in case study research. *Industrial Marketing Management,* 39**,** 118-128.

EASTWOOD, J. G., JALALUDIN, B. B. & KEMP, L. A. 2014. Realist explanatory theory building method for social epidemiology: a protocol for a mixed method multilevel study of neighbourhood context and postnatal depression. *Springerplus,* 3**,** 12-12.

EBENSO, B., MANZANO, A., UZOCHUKWU, B., ETIABA, E., HUSS, R., ENSOR, T., NEWELL, J., ONWUJEKWE, O., EZUMAH, N., HICKS, J. & MIRZOEV, T. Dealing with context in logic model development: Reflections from a realist evaluation of a community health worker programme in Nigeria. 2019 2019 Great Britain. Elsevier Science B.V., Amsterdam., 97.

EDGLEY, A., STICKLEY, T., TIMMONS, S. & MEAL, A. Critical realist review: exploring the real, beyond the empirical. 2016 2016 Great Britain. Taylor & Francis, 316.

EINARSON, T., ACS, A., LUDWIG, C. & PANTON, U. 2018. Prevalence of cardiovascular disease in type 2 diabetes: a systematic literature review of scientific evidence from across the world in 2007–2017. *Cardiovascular Diabetology,* 17.

EL HUSSEIN, M., SALYERS, V. & OSUJI, J. 2019. Measuring the Effectiveness of Visual Narrative Illustrations for Learning Pathophysiology Concepts. *Transformative Dialogues: Teaching & Learning Journal,* 12**,** 1.

ELKOUBAITI, H. & MRABET, R. 2019. Key Elements of Educational Augmented and Virtual Reality Applications. *Information Systems and Technologies to Support Learning,* 111.

ELLIS, R. D. & NEWTON, N. 2005. *Consciousness & emotion: Agency, conscious choice, and selective perception,* Amsterdam, John Benjamins Publishing Company.

ENGLUND, C., GUSTAFFSON, M. & GALLEGO, G. 2017. Pharmacy Students' Attitudes and Perceptions of "Virtual Worlds" as an Instructional Tool for Clinical Pharmacy Teaching. *Pharmacy*.

ENYEDY, N., DANISH, J., DELACRUZ, G. & KUMAR, M. 2012. Learning physics through play in an augmented reality environment. *International Journal of Computer-Supported Collaborative Learning,* 7**,** 347-378.

FALCONER, L. 2013. Situated learning in accident investigation: a virtual world simulation case study. *International Journal of Learning Technology,,* 8.

FALCONER, L. 2017. Experiencing sense of place in virtual and physical Avebury . *Personal and Ubiquitous Computing,* 21**,** 977-988.

FALCONER, L., CARMEN, M. & ORTEGA, G. 2018. *Virtual Worlds: Concepts, Applications and Future Directions,* USA, Nova Science Publishers,.

FARROW, R. & IACOVIDES, I. 2014. Gaming and the limits of digital embodiment. Netherlands: Springer.

FEALY, S., DONOVAN, J., HUTTON, A., GRAHAM, K., MCNEILL, L., SWEET, L. & HAZELTON, M. 2019. The integration of immersive virtual reality in tertiary nursing and midwifery education: A scoping review. *Nurse Education Today,* 79**,** 14-19.

FERNANDEZ, M. 2017. Augmented virtual reality: how to improve education systems. *Higher Learning Research Communities,* 7**,** 1-15.

FINCHAM, J. 2008. Response Rates and Responsiveness for Surveys, Standards, and the Journal. *American Journal of Pharmaceutical Education,* 72**,** 1-3.

FLETCHER, A. J. Applying critical realism in qualitative research: methodology meets method. 2017 2017 Great Britain. Taylor & Francis, 181.

FORD, J., JONES, A., WONG, G., CLARK, A., PORTER, T. & STEEL, N. 2018. Access to primary care for socio-economically disadvantaged older people in rural areas: exploring realist theory using structural equation modelling in a linked dataset. *BMC Medical Research Methodology***,** 1.

FORONDA, C. L., SWOBODA, S. M., HUDSON, K. W., JONES, E., SULLIVAN, N., OCKIMEY, J. & JEFFRIES, P. R. 2016. Featured Article: Evaluation of vSIM for Nursing™: A Trial of Innovation. *Clinical Simulation in Nursing,* 12**,** 128-131.

FOWLER, C. 2015. Virtual reality and learning: where is the pedagogy? *BJET,* 46**,** 412-422.

FREDERICKS, S., SIDANI, S., FOX, M. & MIRANDA, J. 2019. Strategies for balancing internal and external validity in evaluations of interventions. *Nurse Researcher,* 27**,** 19-23.

FREINA, L. & OTT, M. 2015. A LITERATURE REVIEW ON IMMERSIVE VIRTUAL REALITY IN EDUCATION: STATE OF THE ART AND PERSPECTIVES. *eLearning & Software for Education***,** 133-141.

FURIÓ, D., GONZÁLEZ-GANCEDO, S., JUAN, M. C., SEGUÍ, I. & RANDO, N. 2013. Evaluation of learning outcomes using an educational iPhone game vs. traditional game. *Computers & Education,* 64**,** 1-23.

GAMAGE, V., TRETIAKOV, T. & CRUMP, B. 2011. Teacher perceptions of learning affordances of multi-user virtual environments. *Computers and Education,* 57**,** 2406-2413.

GARZON, J. & ACEVEDO, J. 2019. Meta-analysis of the impact of Augmented Reality on students' learning gains. *Educational Research Review,* 27**,** 244-260.

GENTRY, S. 2020. Serious Gaming and Gamification interventions for health professional education. *Cochrane Database of Systematic Reviews*.

GERRISH, K. & LATHLEAN, J. 2015. *The Research process in nursing.,* Gerrish K, Lathlean J. Chichester John Wiley and Sons Ltd.

GHANBARZADEH, R. & GHAPANCHI, A. 2018. Investigating various application areas of three-dimensional virtual worlds for higher education. *BJET,* 49**,** 370-384.

GILL, M., ANDERSEN, E. & HILSMANN, N. Best practices for teaching pharmacology to undergraduate nursing students: A systematic review of the literature. 2019 2019 Great Britain. Elsevier Science B.V., Amsterdam, 15.

GILMORE, B., ADAMS, B. J., BARTOLONI, A., ALHAYDAR, B., MCAULIFFE, E., RAVEN, J., TAEGTMEYER, M. & VALLIÈRES, F. 2016. Improving the performance of community health workers in humanitarian emergencies: a realist evaluation protocol for the PIECES programme. *BMJ Open,* 6**,** e011753-e011753.

GIRVAN, C. & SAVAGE, T. Virtual worlds: A new environment for constructionist learning. 2019 2019 Great Britain. Elsevier Science B.V., Amsterdam., 396.

GIRVAN, C., TANGNEY, B. & SAVAGE, T. 2013. SLurtles: Supporting constructionist learning in Second Life. *Computers & Education,* 61**,** 115-132.

GODWIN-JONES 2016. AUGMENTED REALITY AND LANGUAGE LEARNING: FROM ANNOTATED VOCABULARY TO PLACE-BASED MOBILE GAMES. *Language Learning & Technology,* 20**,** 9-19.

GOESER, P. T., JOHNSON, W. M., HAMZA-LUP, F. G. & SCHAEFER, D. 2011. VIEW - A Virtual Interactive Web-based Learning Environment for Engineering. *Advances in Engineering Education,* 2**,** 1.

GOOCH, D., VASALOU, A., BENTON, L. & KHALED, R. 2016. Using Gamification to Motivate Students with Dyslexia. *Conference on Human Factors in Computing Systems Proceedings***,** 969.

GOSMANOV, A. R. 2018. Use of Inpatient Glucose Values in Identifying Patients at High Risk for Diabetes. *American Journal of Medicine,* 131**,** e395-e395.

GOUGH, D., OLIVER, S. & THOMSON, J. 2012. *An Introduction to Systematic Reviews,* Sage Publications, London.

GREENWAY, K., BUTT, G. & WALTHALL, H. What is a theory-practice gap? An exploration of the concept. 2019 2019 Great Britain. ELSEVIER SCIENCE B.V. AMSTERDAM, 1.

GUSTAFSSON, M., GALLEGO, G. & ENGLUND, C. 2017. The description and evaluation of virtual worlds in clinical pharmacy education in Northern Sweden. *Currents in Pharmacy Teaching & Learning,* 9**,** 887-892.

HACHAJ, T. & BARANIEWICZ, D. 2015. Knowledge Bricks—Educational immersive reality environment. *International Journal of Information Management,* 35**,** 396-406.

HACK, C. J. 2016. The benefits and barriers of using virtual worlds to engage healthcare professionals on distance learning programmes. *Interactive Learning Environments,* 24**,** 1836-1849.

HAIR, J., HULT, G., RINGLE, C. & SARSTEDT, M. 2016. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM),* London, Sage.

HAIR, J., HULT, G., RINGLE, C., SARSTEDT, M. & THIELE, K. 2017. Mirror, mirror on the wall: a comparative evaluation of composite-based structural equation modeling methods. *Journal of the Academy of Marketing Science,* 45**,** 616.

HAIR, J., SARSTEDT, M., HOPKINS, L. & VOLKER, G. 2014. Partial least squares structural equation modeling (PLS-SEM) : An emerging tool in business research. *European Business Review,* 26**,** 106-121.

HAIR, J., SARSTEDT, M., RINGLE, C. & GUDERGAN, S. 2018. *Advanced Issues in Partial Least Squares Structural Equation Modeling,* Sage, London.

HAIR, J. F., SARSTEDT, M., PIEPER, T. M. & RINGLE, C. M. 2012. The Use of Partial Least Squares Structural Equation Modeling in Strategic Management Research: A Review of Past Practices and Recommendations for Future Applications. *Long Range Planning,* 45**,** 320-340.

HALL, J. & SAMMONS, P. 2013. Mediation, Moderation & Interaction: Definitions, Discrimination & (Some) Means of Testing. *In:* TEO, T.

(ed.) *Handbook of Quantitative. Methods for Educational Research. .* Rotterdam.: Sense Publishers.

HANKS, E. & ECKSTEIN, G. 2019. Increasing English Learners' Positive Emotional Response to Learning Through Dance. *TESL Reporter,* 52**,** 72-93.

HARDING, J., PAVKOV, M., MAGLIANO, D., SHAW, J. & GREG, E. 2019. Global trends in diabetes complications: a review of current evidence. *Diabetologia,* 62**,** 3-16.

HASE, S. S. H. S. E. A. & KENYON, C. 2007. Heutagogy: A Child of Complexity Theory. *Complicity: An International Journal of Complexity & Education,* 4**,** 111-117.

HATTIE, J. 2015. The applicability of Visible Learning to higher education. *Scholarship of Teaching and Learning in Psychology,* 1**,** 79-91.

HAUZE, S. W., HOYT, H. H., FRAZEE, J. P., GREINER, P. A. & MARSHALL, J. M. 2019. Enhancing Nursing Education Through Affordable and Realistic Holographic Mixed Reality: The Virtual Standardized Patient for Clinical Simulation. United States: Kluwer Academic Publishers.

HEE LEE, J. & SHVETSOVA, O. 2019. The Impact of VR Application on Student’s Competency Development: A Comparative Study of Regular and VR Engineering Classes with Similar Competency Scopes. *Sustainability,* 11**,** 2221.

HENSELER, J. & HUBONA, G. 2016. Using PLS path modeling in new technology research: updated guidelines. *Industrial Management & Data Systems,* 116**,** 2-20.

HEW, K. & CHEUNG, W. 2013. Review: Use of Web 2.0 technologies in K-12 and higher education: The search for evidence-based practice. *Educational Research Review,* 9**,** 47-64.

HIGGINS, A., O'HALLORAN, P. & PORTER, S. 2012. Management of Long Term Sickness Absence: A Systematic Realist Review. *Journal of Occupational Rehabilitation,* 22**,** 322-332.

HIGGINS, J. & GREEN, S. 2011. *Cochrane Handbook for Systematic Reviews of Interventions: The Cochrane Collaboration,* Chichester., Wiley-Blackwell,.

HILL, B. 2017. Research into experiential learning in nurse education. *British Journal of Nursing,* 26**,** 932-938.

HIRT, J. & BEER, T. 2020. Use and impact of virtual reality simulation in dementia care education: A scoping review. *Nurse Educatio Today,* 84.

HJØRLAND, B. & WIKGREN, M. 2005. Critical realism as a philosophy and social theory in information science? *Journal of Documentation,* 61**,** 11-22.

HOBBS, M. & HOLLEY, D. Using augmented reality to engage STEM students with an authentic curriculum. 2016 In: Vincenti G., Bucciero A., Vaz de Carvalho C. (eds) E-Learning, E-Education, and Online Training. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, . Springer, Cham.

HOFER, M., HÜSSER, A. & PRABHU, S. 2017. The effect of an avatar's emotional expressions on players' fear reactions: The mediating role of embodiment. *Computers in Human Behavior,* 75**,** 883-890.

HOFFMANN, H., STEFANI, O. & PATEL, H. 2006. Extending the desktop workplace by a portable virtual reality system. Great Britain: Elsevier Science B.V., Amsterdam.

HOLDER, A. & BETHEA-HAMPTON, T. The Results of Using Open Educational Resources and Virtual Reality in Higher Education. Annual Meeting of the Association Supporting Computer Users in Education, 2019 North Carolina. Association Supporting Computer Users in Education.

HOLLEY, D. & HOBBS, M. 2020. *Augmented Reality For Education*, Springer Nature.

HORN, J. L. & MCARDLE, J. J. 1992. A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research,* 18**,** 117-144.

HORWOOD, C. R., RAYO, M. F., FITZGERALD, M., BALKIN, E. A. & MOFFATT-BRUCE, S. D. 2018. Gaps Between Alarm Capabilities and Decision-making Needs: An Observational Study of Detecting Patient Decompensation. *Proceedings of the International Symposium of Human Factors & Ergonomics in Healthcare,* 7**,** 112.

HOUSTON, S. & MONTGOMERY, L. 2017. Reflecting critically on contemporary social pathologies: social work and the 'good life'. *Critical and Radical Social Work,* 5**,** 181.

HUANG, G. Q., YANG, T. H. & XU, S. 2013. Application of Virtual Reality Technology in Teaching. *Applied Mechanics and Materials,* 475**,** 1230.

HUANG, H. M., RAUCH, U. & LIAW, S. S. 2010. Investigating Learners' Attitudes toward Virtual Reality Learning Environments: Based on a Constructivist Approach. *Computers and Education,* 55**,** 1171-1182.

HUANG, T.-C., CHEN, C.-C. & CHOU, Y.-W. 2016. Animating eco-education: To see, feel, and discover in an augmented reality-based experiential learning environment. *Computers & Education,* 96**,** 72-82.

HULT, G. T. M. & KETCHEN JR, D. J. 2001. DOES MARKET ORIENTATION MATTER?: A TEST OF THE RELATIONSHIP BETWEEN POSITIONAL ADVANTAGE AND PERFORMANCE. *Strategic Management Journal (John Wiley & Sons, Inc.),* 22**,** 899.

HWANG, W. Y. & HU, S. S. 2013. Analysis of peer learning behaviours using multiple representations in virtual reality and their impact on geometry problem solving. *Computers & Education,* 62**,** 308-319.

INMAN, C., WRIGHT, V. & HARTMAN, J. 2010. Use of Second Life in K-12 and Higher Education: A Review ofResearch. *Journal of Interactive Online Learning,* 9**,** 44-63.

INMAN, C., WRIGHT, V. & HARTMAN, J. 2011. USE OF SECOND LIFE IN K-12 AND HIGHER EDUCATION: A Review of Research. *The Turkish Online Journal of Distance Education, Vol 12, Iss 3/2, Pp 67-85 (2011)***,** 67.

IRWIN, P. & COUTTS, R. 2015. A Systematic Review of the Experience of Using Second Life in the Education of Undergraduate Nurses. *Journal of Nursing Education,* 54**,** 572-577.

ISHAQ, O., ADISA, A., L., V., P., R. & R., J. 2019. Recent advances in nano-enabled fertilizers and pesticides: a critical review of mechanisms of action. *Environmental Science: Nano*.

İSMAILOĞLU, E. G. & ZAYBAK, A. 2018. Comparison of the Effectiveness of a Virtual Simulator With a Plastic Arm Model in Teaching Intravenous Catheter Insertion Skills. *CIN: Computers, Informatics, Nursing,* 36**,** 98.

JAGOSH, J., PLUYE, P., WONG, G., CARGO, M., SALSBERG, J., BUSH, P. L., HERBERT, C. P., GREEN, L. W., GREENHALGH, T. & MACAULAY, A. C. 2014. Critical Reflections on Realist Review: Insights from Customizing the Methodology to the Needs of Participatory Research Assessment. *Research Synthesis Methods,* 5**,** 131-141.

JANSSEN, D., TUMMEL, C., RICHERT, A. & ISENHARDT, I. 2016. Virtual Environments in Higher Education - Immersion as a Key Construct for Learning 4.0. *International Journal of Advanced Corporate Learning,* 9**,** 20-26.

JÄRVENSIVU, T. & TÖRNROOS, J.-Å. 2010. Case study research with moderate constructionism: Conceptualization and practical illustration. *Industrial Marketing Management,* 39**,** 100-108.

JEE, M., J. 2014. From First Life to Second Life: Evaluating task-based language learning in a new environment / De la vie réelle à la vie virtuelle: évaluation de l'apprentissage des langues basé sur les tâches dans un nouvel environnement. *Canadian Journal of Learning and Technology, Vol 40, Iss 1 (2014)*.

JOHNSON, C., HILL, L., LOCK, J., ALTOWAIRIKI, N., OSTROWSKI, C., DA ROSA DOS SANTOS, L. & LIU, Y. 2017. Using Design-Based Research to Develop Meaningful Online Discussions in Undergraduate Field Experience Courses. *International Review of Research in Open and Distributed Learning,* 18**,** 36-53.

JUDEN, M. 2014. *Realist randomised controlled trials of development interventions in practice: concrete design suggestions to address the problem of external validity.* London.

JURAS, G., BRACHMAN, A., MICHALSKA, J., KAMIENIARZ, A., PAWLOWSKI, M., HADAMUS, A., BIALOSZEWSKI, D., BLASZCYK, J. & SLOMA, K. 2019. Standards of Virtual Reality Application in Balance Training Programs in Clinical Practice: A Systematic Review. *Games for Health Journal,* 8.

JUSTICE, C., RICE, J., WARRY, W., INGLIS, S., MILLER, S. & SAMMON, S. 2007. Inquiry in Higher Education: Reflections and Directions on Course Design and Teaching Methods. *Innovative Higher Education,* 31**,** 201-214.

KAMARAINEN, A. M., METCALF, S., GROTZER, T., BROWNE, A., MAZZUCA, D., TUTWILER, M. S. & DEDE, C. 2013. EcoMOBILE: Integrating augmented reality and probeware with environmental education field trips. *Computers & Education,* 68**,** 545-556.

KARDONG-EDGREN, S., FARRA, S. L., ALINIER, G. & YOUNG, H. M. A Call to Unify Definitions of Virtual Reality. 2019 2019 Netherlands. ELSEVIER SCIENCE B.V. AMSTERDAM, 28.

KARKAR, S., GLOAGUEN, A., LE GUEN, Y., PIERRE-JEAN, M., DANDINE-ROULLAND, C., LE FLOCH, E., PHILIPPE, C., TENENHAUS, A. & FROUIN, V. 2019. Multivariate Haplotype Analysis Of 96 Sulci Opening For 15,612 UK-Biobank Sujects. IEEE.

KAWULICH, B. & D'ALBA, A. 2018. Teaching qualitative research methods with Second Life, a 3-dimensional online virtual environment. *Virtual reality***,** 1-10.

KE, F. & HSU, Y.-C. 2015. Mobile augmented-reality artifact creation as a component of mobile computer-supported collaborative learning. *The Internet and Higher Education,* 26**,** 33-41.

KEIJSERS, C. J., BROUWERS, J. R., WILDT, D. J., CUSTERS, E. J., TEN CATE, O. T., HAZEN, A. C. & JANSEN, P. A. 2014. A comparison of medical and pharmacy students' knowledge and skills of pharmacology and pharmacotherapy. Great Britain: Blackwell Publishing Ltd.

KELLY, K. & LANIER, J. 1989. *The "Real" Worlds of Virtual Space,* UK, Springer.

KERAWALLA, L., LUCKIN, R., SELJEFLOT, S. & WOOLARD, A. 2006. “Making it real”: exploring the potential of augmented reality for teaching primary school science. *Virtual Reality,* 10**,** 163.

KERR, H., PRICE, J., NICHOLL, H. & O'HALLORAN, P. Transition from children's to adult services for young adults with life-limiting conditions: A realist review of the literature. 2017 2017 Great Britain. Elsevier Science B.V., Amsterdam., 1.

KESKITALO, T. 2011. Teachers' conceptions and their approaches to teaching in virtual reality and simulation-based learning environments. *Teachers and Teaching: Theory and Practice,* 17**,** 131-147.

KESKITALO, T. & RUOKAMO, H. 2015. A PEDAGOGICAL MODEL FOR SIMULATION-BASED LEARNING IN HEALTHCARE. *Seminar.Net: Media, Technology & Life-Long Learning,* 11**,** 1-1.

KIDD, L., KNISLEY, S. & MORGAN, K. 2012. Effectiveness of a Second Life® Simulation as a Teaching Strategy for Undergraduate Mental Health Nursing Students. *Journal of Psychosocial Nursing and Mental Health Services,* 50**,** 28-37.

KIILI, K., DE FREITAS, S., ARNAB, S. & LAINEMA, T. 2012. The Design Principles for Flow Experience in Educational Games. *Procedia Computer Science,* 15**,** 78-91.

KIMMONS, R. & JOHNSTUN, K. 2019. Navigating Paradigms in Educational Technology. *TechTrends: Linking Research and Practice to Improve Learning,* 63**,** 631-641.

KING, D., TEE, S., FALCONER, L., ANGELL, C., HOLLEY, D. & MILLS, A. 2018. Virtual health education: Scaling practice to transform student learning: Using virtual reality learning environments in healthcare education to bridge the theory/practice gap and improve patient safety. *Nurse Education Today,* 71**,** 7-9.

KIRKMAN, M. A., AHMED, M., ALBERT, A. F., WILSON, M. H., NANDI, D. & SEVDALIS, N. 2014. The use of simulation in neurosurgical education and training. A systematic review. *Journal Of Neurosurgery,* 121**,** 228-246.

KIRRIEMUIR, J. 2008. Second Life in higher education, medicine and health. *BJET,* 64.

KITCHENHAM, B. 2004. Procedures for performing systematic reviews. . In Keele University Technical Report TR/SE-0401 and NICTA Technical Report 0400011T.1 . Australia.

KLEVJER, R. 2008. Diane Carr, David Buckingham, Andrew Burn and Gareth Scott, Computer Games: Text, Narrative and Play. Great Britain: SAGE PUBLICATIONS.

KLOPFER, E. & SQUIRE, K. 2008. Environmental Detectives—The Development of an Augmented Reality Platform for Environmental Simulations. *Educational Technology Research and Development***,** 203.

KONSTANTINIDIS, A., THRASYVOULOS, T., THEODOULI, T. & ANDREAS, P. 2010. Fostering collaborative learning in Second Life: Metaphors and affordances. *Computers and Education,* 55**,** 603-615.

KRISTIANSEN, T. & GRØNKJÆR, M. 2017. Focus Groups as Social Arenas for the Negotiation of Normativity. *International Journal of Qualitative Methods*.

KUNST, E., HENDERSON, A. & JOHNSTON, A. 2018. A Scoping Review of the Use and Contribution of Simulation in Australian Undergraduate Nurse Education. *Clinical Simulation in Nursing,* 19**,** 17-29.

KURILOVAS, E. 2016. Evaluation of quality and personalisation of VR/AR/MR learning systems. *Behaviour & Information Technology,* 35**,** 998-1007.

KURILOVAS, E., DVARECKIENĖ, V. & JEVSIKOVA, T. 2016. Augmented Reality-Based Learning Systems: Personalisation Framework. *Proceedings of the European Conference on e-Learning***,** 391.

LAU, K. & LEE, P. 2015. The use of virtual reality for creating unusual environmental stimulation to motivate students to explore creative ideas. *Interactive Learning Environments,* 23**,** 3-18.

LAZAROS, M., EMILI, B.-B. & XUN, H. 2018. Flow and Immersion in Video Games: The Aftermath of a Conceptual Challenge. *Frontiers in Psychology*.

LEE, A., WONG, K. & FUNG, C. 2010. How does desktop virtual reality enhance learning outcomes? A structural equation modeling approach. *Computers & Education,* 55**,** 1424-1442.

LEE, B.-O., LIANG, H.-F., CHU, T.-P. & HUNG, C.-C. 2019. Effects of simulation-based learning on nursing student competences and clinical performance. *Nurse Education in Practice,* 41.

LEE, K. 2012. Augmented Reality in Education and Training. *TechTrends: Linking Research & Practice to Improve Learning,* 56**,** 13-21.

LEIGH, J. & ROBERTS, D. 2018. Critical exploration of the new NMC standards of proficiency for registered nurses. *British Journal of Nursing,* 27**,** 1068-1072.

LIAROKAPIS, F., KOUŘIL, P., AGRAFIOTIS, P., DEMESTICHA, S., CHMELÍK, J. & SKARLATOS, D. 2017. 3D MODELLING AND MAPPING FOR VIRTUAL EXPLORATION OF UNDERWATER ARCHAEOLOGY ASSETS. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol XLII-2/W3, Pp 425-431 (2017)***,** 425.

LIAW, S. S., SOH, S., TAN, K., WU, L., YAP, J., CHOW, Y., LAU, T. & LIM, W. Design and evaluation of a 3D virtual environment for collaborative learning in interprofessional team care delivery. 2019 2019 Great Britain. Elsevier Science B.V., Amsterdam, 64.

LIMNIOU, M., ROBERTS, D. & PAPADOPOULOS, N. 2008. Full Immersive Virtual Environment Cave[TM] in Chemistry Education. *Computers & Education,* 51**,** 584-593.

LINDGREN, R. & JOHNSON-GLENBERG, M. 2013. Emboldened by Embodiment: Six Precepts for Research on Embodied Learning and Mixed Reality. *Educational Researcher,* 42**,** 445-452.

LINSLEY, P., HOWARD, D. & OWEN, S. 2015. The construction of context-mechanisms-outcomes in realistic evaluation. *Nurse Researcher,* 22**,** 28-34.

LIU, T.-Y. & CHU, Y.-L. 2010. Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education,* 55**,** 630-643.

LOHRE, R., BOIS, A. J., ATHWAL, G. S. & GOEL, D. P. 2020. Improved Complex Skill Acquisition by Immersive Virtual Reality Training: A Randomized Controlled Trial. *The Journal Of Bone And Joint Surgery. American Volume*.

LONGO, M. R., SCHÜÜR, F., KAMMERS, M. P. M., TSAKIRIS, M. & HAGGARD, P. 2008. What is embodiment? A psychometric approach. *Cognition,* 107**,** 978-998.

LOPREIATO, J., DOWNING, D., GAMMON, W., LIOCE, L., SITNER, B. & SLOT, V. 2016. Healthcare Simulation Dictionary. Healthc Simul Dict 2016.

LOWRY, P. B. & GASKIN, J. 2014. Partial Least Squares (PLS) Structural Equation Modeling (SEM) for Building and Testing Behavioral Causal Theory: When to Choose It and How to Use It. *IEEE Transactions on Professional Communication, Professional Communication, IEEE Transactions on, IEEE Trans. Profess. Commun.,* 57**,** 123-146.

LØYLAND, B., ANGELHOFF, C., KRISTJÁNSDÓTTIR, G. & SJØLIE, H. 2020. A systematic integrative review of parents' experience and perception of sleep when they stay overnight in the hospital together with their sick children. *Journal of Clinical Nursing (John Wiley & Sons, Inc.),* 29**,** 706-719.

MACLURE, M. 2007. 'Clarity bordering on stupidity': where's the quality in systematic review? *Journal of Educational Policy,* 20**,** 393-416.

MAHON, J., BRYANT, B., BROWN, B. & KIM, M. 2010. Using Second Life to enhance classroom management practice in teacher education. *Educational Media International,* 47**,** 121-134.

MAKRANSKY, G., TERKILDSEN, T. S. & MAYER, R. E. Adding immersive virtual reality to a science lab simulation causes more presence but less learning. 2019 2019 Great Britain. Elsevier Science B.V., Amsterdam., 225.

MARAVILLA, M., CISNEROS, A., STODDARD, A., SCRETCHING, D., MURRAY, B. & REDMILES, E. 2019. Defining virtual reality: Insights from research and practice. *iConference 2019 Proceedings*.

MARCEL, F. 2019. Mobile augmented reality learning objects in higher education. *Research in Learning Technology,* 27.

MARCHAL, B., VAN BELLE, S., VAN OLMEN, J., HOEREE, T. & KEGELS, G. 2012. Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. Great Britain: SAGE PUBLICATIONS.

MARKOVA, I., GROSSEN, M. & LINELL, P. 2007. *Dialogue in Focus Groups : Exploring Socially Shared Knowledge,* London, Equinox Publishing Ltd.

MARTIN-GONZALEZ, A., CHI-POOT, A. & UC-CETINA, V. 2016. Usability evaluation of an augmented reality system for teaching Euclidean vectors. *Innovations in Education & Teaching International,* 53**,** 627-636.

MARTIN-GUTIERREZ, J., FERNANDEZ, M. & DOLORES, M. 2014. Applying Augmented Reality in Engineering Education to Improve Academic Performance & Student Motivation. *International Journal of Engineering Education,* 30**,** 625.

MARTIN, O. & CARR, D. 2009. Learning in virtual worlds: Using communities of practice to explain how people learn from play. *Part of the special issue on Multi-user virtual environments,* 40**,** 444-457.

MARTIN, S., VALLANCE, M., VAN SCHAIK, P. & WIZ, C. 2010. Learning spaces, tasks and metrics for effective communication in Second Life within the context of programming LEGO NXT Mindstorms™ robots: towards a framework for design and implementation. *Journal of Virtual Worlds Research,* 3**,** 3-24.

MATHEWS, S., ANDREWS, L. & LUCK, E. 2012. Developing a Second Life virtual field trip for university students: an action research approach. *Educational Research,* 54**,** 17-38.

MATT, E. 2018. The Impact of Embedded iPad Use on Student Perceptions of Their Digital Capabilities. *Education Sciences,* 8.

MATTAR, J. 2018. Constructivism and connectivism in education technology: Active, situated, authentic, experiential, and anchored learning. *Revista de Educaron a Distancia,* 21.

MATTHEW, A. & BUTLER, D. 2017. Narrative, machinima and cognitive realism: Constructing anauthentic real-world learning experience for law students. *Australasian Journal of Educational Technology,* 33**,** 48-162.

MATTHEWS, L., HAIR, J. O. E. & MATTHEWS, R. 2018. PLS-SEM: THE HOLY GRAIL FOR ADVANCED ANALYSIS. *Marketing Management Journal,* 28**,** 1-13.

MAYER, R. E., GRIFFITH, E., JURKOWITZ, I. T. N. & ROTHMAN, D. 2008. Increased Interestingness of Extraneous Details in a Multimedia Science Presentation Leads to Decreased Learning. United States: AMERICAN PSYCHOLOGICAL ASSOCIATION.

MAYRATH, M. C., TRAPHAGAN, T., JARMON, L., TRIVEDI, A. & RESTA, P. 2010. Teaching with virtual worlds: factors to consider for instructional use of Second Life. *Journal of Educational Computing Research,* 43**,** 403-444.

MCCONNELL, T. & PORTER, S. 2017. Music therapy for palliative care: A realist review. *Palliative & Supportive Care,* 15**,** 454-464.

MCEVOY, P. & RICHARDS, D. 2006. A critical realist rationale for using a combination of quantitative and qualitative methods. *Journal of Research in Nursing,* 11**,** 66-78.

MCFAUL, H. & FITZGERALD, E. 2019. A realist evaluation of student use of a virtual reality smartphone application in undergraduate legal education. *British Journal of Educational Technology*.

MEGALI, G., SINIGAGLIA, S., TONET, O. & DARIO, P. 2006. Modelling and Evaluation of Surgical Performance Using Hidden Markov Models. *IEEE Transactions on Biomedical Engineering, Biomedical Engineering, IEEE Transactions on, IEEE Trans. Biomed. Eng.***,** 1911.

MERCHANT, Z., GOETZ, E., CIFUENTES, L., KEENEY-KENNICUTT, W. & DAVIS, T. 2014. Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education,* 70**,** 29-40.

MIETTINEN, M. 2019. *The Learning Curve of Reality.* Sydney.

MIKROPOULOS, T. & NATSIS, A. 2011. Educational virtual environments: A ten-year review of empirical research. *Computers & Education***,** 769-780.

MILK, C. 2015. How virtual reality can create the ultimate empathy machine.

MILLER, R. 2014. The Application of Virtual Reality in Higher Education Distance Learning. *Journal of Applied Learning Technology,* 4**,** 15.

MINGERS, J. 2004. Real-izing information systems: critical realism as an underpinning philosophy for information systems. Great Britain: Elsevier Science B.V., Amsterdam.

MINOCHA, S. & REEVES, A. J. 2010. Design of learning spaces in 3D virtual worlds: an empirical investigation of Second Life. *Learning, Media & Technology,* 35**,** 111-137.

MKANDAWIRE-VALHMU, L., WENDLAND, C., STEVENS, P. E., KAKO, P. M., DRESSEL, A. & KIBICHO, J. 2013. Marriage as a risk factor for HIV: Learning from the experiences of HIV-infected women in Malawi. Great Britain: Taylor & Francis.

MOGLIA, A., FERRARI, V., MORELLI, L., FERRARI, M., MOSCA, F. & CUSCHIERI, A. 2016. Platinum Priority – Review – Education: A Systematic Review of Virtual Reality Simulators for Robot-assisted Surgery. *European Urology,* 69**,** 1065-1080.

MOHER, D., LIBERATI, A., TETZLAFF, J. & ALTMAN, D. 2009. Preferred reporting items for systematic reviews and meta-analyses: the

PRISMA statement. *PLoS Medicine, Vol 6, Iss 7, p e1000097 (2009)*, e1000097.

MORALES, R., GUEVARA-CANALES, J., FLORES-LUJAN, V., ROBELLO-MALATTO, J. & BAZAN-ASCENCIOS, R. 2019. Use of VR as a learning environment in dentistry. *General Denistry.,* 67, 21-27.

MORGAN, D. 2019. *Basic and advanced focus groups*, SAGE Publications, Inc.

MORLEY, D., BETTLES, S. & DERHAM, C. The exploration of students' learning gain following immersive simulation – the impact of feedback. 2019 2019. Taylor & Francis, 368.

MORRISON, B. B., DORN, B. & GUZDIAL, M. 2014. Measuring cognitive load in introductory CS. *International Computing Education Research Workshop*, 131.

MUNNERLEY, D., BACON, M., FITZGERALD, R., WILSON, A., HEDBERG, J. & STEELE, J. 2014. Augmented Reality: application in Higher Education. Australian Government Office for Learning and Teaching.

MUYLAERT, C., SARUBBI, V., GALLO, P. & NETO, M. 2014. Narrative interviews: an important resource in qualitative research. *Revista da Escola de Enfermagem da USP*, 184.

NEWMAN, S. 2020. Learn How We Learn. *TD: Talent Development,* 74, 12-13.

NITZL, C. 2016. The use of partial least squares structural equation modelling (PLS-SEM) in management accounting research: Directions for future theory development. *Journal of Accounting Literature,* 37, 19-35.

NMC 2018. Nursing and Midwifery Council. Future nurse: standards of proficiency for registered nurses. London.

NUNEZ, D. & BLAKE, E. 2001. Cognitive presence as a unified concept of virtual reality effectiveness. *Proceedings of the 1st International Conference: Computer Graphics, Virtual Reality & Visualisation*, 115.

NUSSLI, N. N. G. C., OH, K. K. U. E. & MCCANDLESS, K. K. D. U. E. 2014. Collaborative Science Learning in Three-Dimensional Immersive Virtual Worlds: Pre-Service Teachers' Experiences in Second Life. *Journal of Educational Multimedia & Hypermedia,* 23, 253-284.

O'ROURKE, P. & CARAMANICA, L. 2019. Addressing Workforce Needs in Skilled Nursing Facilities. *Nurse Leader*.

OGBUANYA, T. C. & ONELE, N. O. Investigating the Effectiveness of Desktop Virtual Reality for Teaching and Learning of Electrical/Electronics Technology in Universities. 2018 2018 United States. Taylor and Francis, 226.

OLEKSY, T. & WNUK, A. 2017. Catch them all and increase your place attachment! The role of location-based augmented reality games in changing people - place relations. *Computers in Human Behaviour,* 76**,** 3-8.

PALINKAS, L., HORWITZ, S., GREEN, C., WISDOM, J., DUAN, N. & HOAGWOOD, K. 2015. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Ad Policy Meant Health,* 42**,** 533-544.

PAMUNGKAS, D. & WARD, K. 2015. Immersive teleoperation of a robot arm using electro-tactile feedback. *2015 6th International Conference on Automation, Robotics & Applications (ICARA)***,** 300.

PARONG, J. & MAYER, R. E. 2018. Learning science in immersive virtual reality. *Journal of Educational Psychology,* 110**,** 785-797.

PARSON, V. & BIGNELL, S. 2017. An Investigation into Cooperative Learning in a Virtual World Using Problem-Based Learning. *Online Learning,* 21**,** 1-16.

PAWSON, R. 2012. Realist thoughts on Cinderella, Alice in Wonderland and health care interventions. *Nursing Inquiry,* 19**,** 4-5.

PAWSON, R. & TILLEY, N. 1997a. An introduction to scientific realist evaluation. *In:* CHELIMSKY, E. & SHADISH, W. R. (eds.) *Evaluation for the 21st century: A handbook.* Thousand Oaks, CA: Sage Publications, Inc.

PAWSON, R. & TILLEY, N. 1997b. *Realistic Evaluation,* Sage Publications.

PELLAS, N. 2016. Bolstering the quality and integrity of online collaborative university- level courses via an open Sim standalone server in conjunction with sloodle. *Education & Information Technologies,* 21**,** 1007-1032.

PENLAND, J. L., LAVIERS, K., BASSHAM, E. & NNOCHIRI, V. 2019. *Virtual Learning: A Study of Virtual Reality for Distance Education." Handbook of Research on Blended Learning Pedagogies and Professional Development in Higher Education.*, IGI Global, .

PEREZ-MARCOS, D., SANCHEZ-VIVES, M. V. & SLATER, M. 2012. Is my hand connected to my body? The impact of body continuity and arm alignment on the virtual hand illusion. Netherlands: Springer Science + Business Media.

PETERSON-AHMAD, M. 2018. Enhancing Pre-Service Special Educator Preparation through Combined Use of Virtual Simulation and Instructional Coaching. *Educational. Sci.,* 8**,** 1-9.

PETTICREW, M. & ROBERTS, H. 2008. *Systematic Reviews in the Social Sciences: A Practical Guide,* Blackwell,, Wiley.

PETTIJOHN, K., PELTIER, C., LUKOS, J., NORRIS, J. & BIGGS, A. 2019. Comparison of Virtual Reality and Augmented Reality: Safety and Effectiveness. USA: Naval Medical Research Unit Dayton Wright-Patterson AFB.

PHILIPS, A., WALZ, A., BERGNER, A., GRAEFF, T., HEISTERMANN, M., IENZLER, S., KORUP, O., LIPP, T., WOLFGANG, S. & ZEILINGER, G. 2015. Immersive 3D geovisualisation in higher education. *Journal of Geography in Higher Education,* 39**,** 437-449.

POLITO, V., BARNIER, A. J., WOODY, E. Z. & CONNORS, M. H. 2014. Measuring agency change across the domain of hypnosis. *Psychology of Consciousness: Theory, Research, and Practice,* 1**,** 3-19.

PORTER, S. 2015. The uncritical realism of realist evaluation. *Evaluation: The International Journal of Theory, Research and Practice,* 21**,** 65-82.

PORTER, S., MCCONNELL, T. & REID, J. 2017. The possibility of critical realist randomised controlled trials. *Trials,* 18**,** 133.

PRESTON, P., LEONE-SHEEHAN, D. & KEYS, B. Nursing student perceptions of pharmacology education and safe medication administration: A qualitative research study. 2019 2019 Great Britain. Elsevier Science B.V., Amsterdam, 76.

PROCTOR, C. Measuring the Computational in Computational Participation: Debugging Interactive Stories in Middle School Computer Science. 13th International Conference on Computer Supported Collaborative Learning, 2019 France. International Society of the Learning Sciences, 104-111.

QUANG, T., AKEEM, P., R., C., HEE TAEK, P., CHAN SIK, P. & HONG KI, K. 2015. A Framework for Using Mobile Based Virtual Reality and Augmented Reality for Experiential Construction Safety Education. *International Journal of Engineering Education,* 31**,** 713.

QUINNEY, A., LAMONT, C., BIGGINS, D. & HOLLEY, D. 2017. Optimising disruptive approaches: extending academic roles and identities in higher education. *Journal of Learning Development in Higher Education,* 12.

RADFORD, M., CONNAWAY, L., CONFER, P., SABOLCSI-BOROS, S. & KWON, H. 2011. "Are We Getting Warmer?" Query Clarification in Live Chat Virtual Reference. *Reference & User Services Quarterly*, 259.

RADIANTI, J., MAJCHRZAK, T. A., FROMM, J. & WOHLGENANNT, I. A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. 2020 2020 Great Britain. Elsevier Science B.V., Amsterdam.

RAIKWAR, A., D'SOUZA , N., ROGERS , C. & KRESS, M. CubeVR: Digital Affordances for Architecture Undergraduate Education using Virtual Reality. 2019 Virtual Reality, IEEE Annual International Symposium.

RAJASEKAR, S., BANGERA, R. K. & SEKARAN, P. 2017. Inter-rater and intra-rater reliability of a movement control test in shoulder. *Journal of Bodywork & Movement Therapies,* 21**,** 739-742.

REEVES, D., HOWELLS, K., SIDAWAY, M., BLAKEMORE, A., HANN, M., PANAGIOTI, M. & BOWER, P. 2018. The cohort multiple randomized controlled trial design was found to be highly susceptible to low statistical power and internal validity biases. *Journal of Clinical Epidemiology,* 95**,** 111-119.

REEVES, T. C. 2005. Design-Based Research in Educational Technology: Progress Made, Challenges Remain. United States: EDUCATIONAL TECHNOLOGY PUBLICATIONS.

RESCHER, N. 2000. Kant on the Limits and Prospects of Philosophy -- Kant, Pragmatism, and the Metaphysics of Virtual Reality. *Kant-Studien,* 91**,** 283.

RIGDON, E. 2012. Rethinking Partial Least Squares Path Modeling: In Praise of Simple Methods. *Long Range Planning,* 45**,** 341-358.

RILEY, J. & KABER, D. 1999. The Effects of Visual Display Type and Navigational Aid on Performance, Presence, and Workload in Virtual Reality Training of Telerover Navigation. United States.

RINGLE, C. & SARSTEDT, M. 2016. Gain more insight from your PLS-SEM results : The importance-performance map analysis. *Industrial Management & Data Systems,* 116**,** 1865-1886.

RINIOLO, T. 1999. Using a large control group for statistical comparison: evaluation of a between-groups median test. *Journal of Experimental Education,* 68**,** 75-88.

RIOU, J., GUYON, H. & FALISSARD, B. 2016. An introduction to the partial least squares approach to structural equation modelling: a method for exploratory psychiatric research. *International Journal of Methods in Psychiatric Research,* 25**,** 220-231.

RITZ, L. & BUSS, A. 2016. A Framework for Aligning Instructional Design Strategies with Affordances of CAVE Immersive Virtual Reality Systems. *TechTrends: Linking Research and Practice to Improve Learning,* 60**,** 549-556.

RIVA, G. 2008. Enacting Interactivity: The Role of Presence. Netherlands: IOS press.

RIZOV, T. & RIZOVA, E. 2015. AUGMENTED REALITY AS A TEACHING TOOL IN HIGHER EDUCATION. *International Journal of Cognitive Research in Science, Engineering & Education (IJCRSEE),* 3**,** 7-16.

ROBINSON, L., ADAIR, P., COFFEY, M., HARRIS, R. & BURNSIDE, G. 2016. Identifying the participant characteristics that predict recruitment and retention of participants to randomised controlled trials involving children: a systematic review. *Trials,* 17**,** 294-294.

ROGERS, L. 2011. Developing Simulations in Multi-User Virtual Environments to Enhance Healthcare Education. *British Journal of Educational Technology,* 42**,** 608-615.

ROSENBAUM, E., KLOPFER, E. & PERRY, J. 2007. On Location Learning: Authentic Applied Science with Networked Augmented Realities. *Journal of Science Education and Technology***,** 31.

ROSS, P. & CROSS, R. 2019. Rise of the e-Nurse: the power of social media in nursing. *Contemporary Nursing,* 55**,** 211-220.

ROTHWELL, P. M. 2006. Factors That Can Affect the External Validity of Randomised Controlled Trials. *PLoS Clinical Trials,* 3**,** 1-5.

ROTKONEN, E., ISLAM, A. & SUTINEN, E. 2019. Toward Pedagogy Driven Virtual Reality Learning Space Design. *Sustainable ICT, Education and Learning***,** 235-244.

ROURKE, S. 2020. How does virtual reality simulation compare to simulated practice in the acquisition of clinical psychomotor skills for pre-registration student nurses? A systematic review. *International Journal of Nursing Studies,* 102.

RYAN, G. S. 2019. Postpositivist critical realism: philosophy, methodology and method for nursing research. *Nurse Researcher,* 27**,** 1-1.

SABAN-BEZALEL, R. & MASHAL, N. 2019. Different Factors Predict Idiom Comprehension in Children and Adolescents with ASD and Typical Development. *Journal of Autism & Developmental Disorders,* 49**,** 4740-4750.

SAIYA, N. 2017. How Dangerous Are Virtual Worlds Really? A Research Note on the Statecraft Simulation Debate. *Social Science Computer Review,* 35**,** 287-296.

SALMON, G., NIE, M. & EDIRISINGHA, P. 2010. Developing a five-stage model of learning in Second Life. *Educational Research,* 52**,** 169-182.

SAMSUDIN, K., RAFI, A., MOHAMAD, A., ZAMZURI, A. & NAZRE, R. 2014. Enhancing a Low-Cost Virtual Reality Application through Constructivist Approach: The Case of Spatial Training of Middle Graders. *Turkish Online Journal of Educational Technology - TOJET,* 13**,** 50-57.

SANDELOWSKI, M. 2001. Real qualitative researchers do not count: the use of numbers in qualitative research. *Research In Nursing & Health,* 24**,** 230-240.

SAPIANO, A., SAMMUT, R. & TRAPANI, J. 2018. The effectiveness of virtual simulation in improving student nurses' knowledge and performance during patient deterioration: A pre and post test design. *Nurse Educaiton Today,* 62**,** 128-133.

SARSTEDT, M. & CHEAH, J.-H. 2019. Partial least squares structural equation modeling using SmartPLS: a software review. *Journal of Marketing Analytics,* 7**,** 196.

SATTAR, M. P., S. LOKMAN, A. HASSAN, A. SHAH, N. RIAZ, Z. 2019. Effects of Virtual Reality training on medical students’ learning motivation and competency. *Pakistan Journal of Medical Sciences,* 35.

SAVIN-BADEN, M. 2013. Spaces in between us: a qualitative study into the impact of spatial practice when learning in Second Life. *London Review of Education,* 11**,** 59-75.

SAVIN-BADEN, M. & FALCONER, L. 2016. Learning at the interstices; locating practical philosophies for understanding physical/virtual inter-spaces. *Interactive Learning Environments,* 24**,** 991-1003.

SAVIN-BADEN, M., GOURLAY, L., TOMBS, C., STEILS, N., TOMBS, G. & MAWER, M. 2010. Situating pedagogies, positions and practices in immersive virtual worlds. *Educational Research,* 52**,** 123-133.

SAVIN-BADEN, M., TOMBS, C., POULTON, T., CONRADI, E., KAVIA, S., BURDEN, D. & BEAUMONT, C. 2011. An evaluation of implementing problem-based learning scenarios in an immersive virtual world. *International Journal of Medical Education,* 2**,** 116-124.

SCHILLER, C. J. 2016. Critical realism in nursing: an emerging approach. *Nursing Philosophy,* 17**,** 88-102.

SCHUBERT, T., FRIEDMANN, F. & REGENBRECHT, H. 2001. The Experience of Presence: Factor Analytic Insights. United States: THE MIT PRESS.

SCHUNK, D. H. 1995. Self-efficacy and education and instruction. *In:* MADDUX, J. E. (ed.) *Self-efficacy, adaptation, and adjustment: Theory, research, and application.* New York, NY: Plenum Press.

SHAFFER, P. Structured causal pluralism in poverty analysis. 2015 2015 Great Britain. Taylor & Francis, 197.

SHMUELI, G. & KOPPIUS, O. R. 2011. PREDICTIVE ANALYTICS IN INFORMATION SYSTEMS RESEARCH. *MIS Quarterly,* 35**,** 553-572.

SIBBALD, B. & ROLAND, M. 1997. Getting research into practice. Great Britain: BLACKWELL SCIENCE.

SIDANI, S., EPSTEIN, D. R. & FOX, M. 2017. Psychometric evaluation of a multi-dimensional measure of satisfaction with behavioral interventions. *Research in Nursing & Health,* 40**,** 459-469.

SIEMENS, G. & CONOLE, G. 2011. Special Issue - Connectivism: Design and Delivery of Social Networked Learning. *International Review of Research in Open and Distance Learning*.

SIMMONDS, H. 2018. Emancipatory marketing and the emancipation of marketing research: a critical realist perspective. *Journal of Critical Realism,* 17**,** 466.

SINGER, M., ALLEN, R., MCDONALD, D. & J., G. 1997. Terrain Appreciation in Virtual Environments: Spatial Knowledge Acquisition.

SINGHAL, S., BAGGA, S., GOYAL, P. & SAXENA, V. 2012. Augmented Chemistry: Interactive Education System. *International Journal of Computer Applications***,** 1-5.

SINGLETON, H., HOLLEY, D., FALCONER, L. & PRIEGO-HERNANDEZ, J. 2019. *Virtual Reality for Nurse Training* [Online]. Available: https://www.bournemouth.ac.uk/why-bu/fusion/using-virtual-reality-train-nurses [Accessed].

SINKOVICS, R., HENSELER, J. & RINGLE, C. 2016. Testing measurement invariance of composites using partial least squares. *International Marketing Review,* 33**,** 405-431.

SLATER, M. 2018. Immersion and the illusion of presence in virtual reality. *British Journal of Psychology,* 109**,** 431-433.

SLATER, M. & WILBUR, S. 1997. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. United States: THE MIT PRESS.

SMITH, K. & KLUMPER, D. 2018. Virtually in the Classroom. *Educational Leadership,* 76**,** 60-65.

SOMMERAUER, P. & MÜLLER, O. 2014. Augmented reality in informal learning environments: A field experiment in a mathematics exhibition. *Computers & Education,* 79**,** 59-68.

SPACEY, A., SCAMMELL, J., BOARD, M. & PORTER, S. 2019. Systematic critical realist review of interventions designed to improve end-of-life care in care homes. *Nursing & Health Sciences*.

SQUIRE, K. 2010. From information to experience: Place-based augmented reality games as a model for learning in a globally networked society. *Teachers College Record,* 112**,** 2565-2602.

SRIVASTAVA, P., RIMZHIM, A., VIJAY, P., SINGH, S. & CHANDRA, S. 2019. Desktop VR Is Better Than Non-ambulatory HMD VR for Spatial Learning. *Frontiers in Robotics and AI*.

STEILS, N., TOMBS, G., MAWER, M., SAVIN-BADEN, M. & WIMPENNY, K. 2015. Implementing the liquid curriculum: the impact of virtual world learning on higher education. *Tech, Pedagogy & Ed,* 24**,** 155-170.

STEIN, C. M., MORRIS, N. J., HALL, N. B. & NOCK, N. L. 2017. Structural Equation Modeling. United States: HUMANA PRESS.

STEVENS, J. & KINCAID, J. P. 2015. The effect of visual display on performance in mixed reality simulation. *International Journal of Modelling & Simulation,* 35**,** 20-26.

STEWART, C. J., STEWART, C., NELSON, A., CAMPBELL, M., WALKER, M., STEVENSON, E., SHAW, J., CUMMINGS, S. & WEST, D. Gut microbiota of Type 1 diabetes patients with good glycaemic control and high physical fitness is similar to people without diabetes: an observational study. Diabetes UK, 2016.

STOCKWELL, B., STOCKWELL, M., CENNAMO, M. & JIANG, E. 2015. *Blended Learning Improves Science Education,* Sciencedirect.com.

STREUKENS, S. & LEROI-WERELDS, S. 2016. Bootstrapping and PLS-SEM: A step-by-step guide to get more out of your bootstrap results. *European Management Journal,* 34**,** 618-632.

SU, C.-H. & CHENG, C.-H. 2013. A Mobile Game-based Insect Learning System for Improving the Learning Achievements. *Procedia - Social and Behavioral Sciences,* 103**,** 42-50.

SU, W., OSISEK, P. & STARNES, B. 2004. Applying the Revised Bloom's Taxonomy to a medical-surgical nursing lesson. *Nurse Educator,* 29**,** 116-120.

SUN SUN, L. & CLARK, L. S. 2010. Virtual worlds as a site of convergence for children's play. *Journal of Virtual Worlds Research,* 3**,** 3-19.

SUNNQVIST, C., KARLSSON, K., LINDELL, L. & FORS, U. 2016. Virtual patient simulation in psychiatric care – A pilot study of digital support for collaborate learning. *Nurse Education in Practice,* 17**,** 30-35.

SWEIGART, L. & HODSON-CARLTON, K. 2013. Improving Student Interview Skills: The Virtual Avatar as Client. *Nurse Educator,* 38**,** 11-15.

TAJIK, E., ABD LATIFF, L., CHIN YIT, S., AWANG, H. & ADZNAM, S. N. A. 2020. Designing and Validating a Questionnaire on Healthy Lifestyle to Reduce Depressive Symptoms among Adolescents. *Iranian Journal of Psychiatry,* 15**,** 27-40.

TAYLOR, C. A. & DUNNE, M. 2011. Virtualization and new geographies of knowledge in higher education: possibilities for the transformation of knowledge, pedagogic relations and learner identities. *British Journal of Sociology of Education,* 32**,** 623-641.

TCHA-TOKEY, K., CHRISTMANN, O., LOUP-ESCANDE, E., LOUP, G. & RICHIR, S. 2018. Towards a Model of User Experience in Immersive Virtual Environments. *Advances in Human-Computer Interaction*.

THOMPSON, C., KIM, R., ALOE, A. & BECKER, B. Extracting the Variance Inflation Factor and Other Multicollinearity Diagnostics from Typical Regression Results. 2017 2017 United States. Taylor & Francis, 81.

THORPE, A. Why has critical realism not been used more in educational leadership and management research? A critical realist exploration. 2020 2020 Great Britain. Maney, 29.

TIFFANY, J. M. & HOGLUND, B. A. 2016. Featured Article: Using Virtual Simulation to Teach Inclusivity: A Case Study. *Clinical Simulation in Nursing,* 12**,** 115-122.

TONDEUR, J., VAN BRAAK, J., SANG, G., VOOGT, J., FISSER, P. & OTTENBREIT-LEFTWICH, A. 2012. Preparing pre-service teachers to integrate technology in education: A synthesis of qualitative evidence. *Computers & Education,* 59**,** 134-144.

TÖRNBERG, P. 2018. Echo chambers and viral misinformation: Modeling fake news as complex contagion. *PLoS ONE,* 13**,** 1.

TRAXLER, J. 2018. Distance Learning--Predictions and Possibilities. *Education Sciences,* 8.

TROCHIM, W., MARCUS, S., MÂSSE, L., MOSER, R. & WELD, P. 2008. The Evaluation of Large Research Initiatives. *American Journal of Evaluation,* 29**,** 8-28.

TURKAN, Y., RADKOWSKI, R., KARABULUT-ILGU, A., BEHZADAN, A. H. & CHEN, A. 2017. Full length article: Mobile augmented reality for teaching structural analysis. *Advanced Engineering Informatics,* 34**,** 90-100.

TÜZÜN, H. & ÖZDINÇ, F. 2016. The effects of 3D multi-user virtual environments on freshmen university students' conceptual and spatial learning and presence in departmental orientation. *Computers & Education,* 94**,** 228-240.

ULRICH, D., FARRA, S., SMITH, S. & HODGOSON, E. 2014. The Student Experience Using Virtual Reality Simulation to Teach Decontamination. *Clinical Simulation in Nursing,* 10**,** 546-553.

VAUGHAN, N., JOHN, N. & REES, N. "ParaVR: Paramedic Virtual Reality Training Simulator," International Conference on Cyberworlds, 2019 Japan. 21-24.

VERKUYL, M., ST-AMANT, O., HUGHES, M., LAPUM, J. L. & MCCULLOCH, T. 2020. Combining Self-Debriefing and Group Debriefing in Simulation. *Clinical Simulation in Nursing,* 39**,** 41-44.

VILLALTA, M., GAJARDO, I., NUSSBAUM, M., ANDREU, J. J., ECHEVERRÍA, A. & PLASS, J. L. 2011. Design guidelines for Classroom Multiplayer Presential Games (CMPG). *Computers & Education,* 57**,** 2039-2053.

VINSON, J. 2019. Be the Change: Understanding Quantitative, Qualitative, & Mixed Methods in Holistic Nursing Research. *Beginnings,* 39**,** 14-16.

WALSH, D. & EVANS, K. 2014. Critical realism: An important theoretical perspective for midwifery research. *Midwifery,* 30**,** e1-e6.

WAN, W. & LAN, A. 2019. The Effectiveness of Virtual Reality-Based Simulation in Health Professions Education Relating to Mental Illness: A Literature Review *Health,* 11**,** 646-660.

WANG, F. & BURTON, J. K. 2013. Second Life in education: A review of publications from its launch to 2011. *British Journal of Educational Technology,* 44**,** 357-371.

WANNER, T. & PALMER, E. Formative self-and peer assessment for improved student learning: the crucial factors of design, teacher participation and feedback. 2018 2018 Great Britain. Taylor & Francis, 1032.

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**,** 414-426.

WARD, T., FALCONER, L., FRUTOS-PEREZ, M., WILLIAMS, B., JOHNS, J. & HAROLD, S. 2016. Using virtual online simulations in Second Life® to engage undergraduate psychology students with employability issues. *British Journal of Educational Technology,* 47**,** 918-931.

WATTS, S. 2018. Effects of Depression, Anxiety, and Stress on Diabetes Biomarkers in African American Adults with Type 2 Diabetes.

WEBSTER, R. **2016**. Declarative knowledge acquisition in immersive virtual learning

Interactive Learning Environments, 24**,** 1319-1333.

WEEKS, K. W., COBEN, D., O'NEILL, D., JONES, A., WEEKS, A., BROWN, M. & PONTIN, D. 2019. Developing and integrating nursing competence through authentic technology-enhanced clinical simulation education: Pedagogies for reconceptualising the theory-practice gap. *Nurse Education in Practice,* 37**,** 29-38.

WEI, X., WENG, D., LIU, Y. & WANG, Y. 2015. Teaching based on augmented reality for a technical creative designcourse. *Computers & Education,* 81**,** 221-234.

WERTS, C. E., ROCK, D. R., LINN, R. L. & JÖRESKOG, K. G. 1978. A General Method of Estimating the Reliability of a Composite. *Educational & Psychological Measurement,* 38**,** 933.

WESTRA, R. 2019. Roy Bhaskar's Critical Realism and the Social Science of Marxian Economics. *Review of Radical Political Economics,* 51**,** 365-382.

WILLIAM, W. & RANDY, J. 1999. Fourteen Propositions About Educational Uses of Virtual Reality. *Educational Technology,* 39**,** 5.

WILLIAMS, L., RYCROFT-MALONE, J. & BURTON, C. R. 2017. Bringing critical realism to nursing practice: Roy Bhaskar's contribution. *Nursing Philosophy,* 18**,** n/a-N.PAG.

WILLIAMS, L. J., VANDENBERG, R. J. & EDWARDS, J. R. 2009. 12 Structural Equation Modeling in Management Research: A Guide for Improved Analysis. *Academy of Management Annals,* 3**,** 543-604.

WIMPENNY, K., SAVIN-BADEN, M., MAWER, M., STEILS, N. & TOMBS, G. 2012. Unpacking frames of reference to inform the design of virtual world learning in higher education. *Australasian Journal of Educational Technology,* 28**,** 522-545.

WITMER, B. G. & SINGER, M. J. 1998. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators & Virtual Environments,* 7**,** 225-240.

WOLD, S., TRYGG, J., BERGLUND, A. & ANTTI, H. 2001. Some recent developments in PLS modeling. *Chemometrics and Intelligent Laboratory Systems,* 58**,** 131-150.

WRIGLEY, T. 2019. The problem of reductionism in educational theory: Complexity, causality, values. *Power and Education*.

WU, H., LEE, S., CHANG, H. & LIANG, J. 2013. Current status, opportunities and challenges of augmented reality in education. *Computers & Education,* 62**,** 41-49.

WYNN, J. D. & WILLIAMS, C. K. 2012. PRINCIPLES FOR CONDUCTING CRITICAL REALIST CASE STUDY RESEARCH IN INFORMATION SYSTEMS. *MIS Quarterly,* 36**,** 787-810.

YADEGARIDEHKORDI, E., NOOR, N., MOHAMAD, A., HANNYZZURA, A. & NORNAZLITA, H. 2019. Affective computing in education: A systematic review and future research. *Computers & Education,* 142.

YI FEI, W., PETRINA, S. & FENG, F. 2017. VILLAGE- Virtual Immersive Language Learning and Gaming Environment: Immersion and presence. *British Journal of Educational Technology,* 48**,** 431-450.

YILDRIM, G., YILDRIM, S. & DOLGUNSOZ, E. 2019. The effect of VR and traditional videos on learner retention and decision making. *World Journal of Ed Tech,* 11.

YILMAZ, Z. & BATDI, V. 2016. A meta-analytic and thematic comparative analysis of the integration of AR applications into education. *Education and Science,* 41**,** 273-289.

YOUNAS, A., PEDERSEN, M. & TAYABEN, J. L. 2019. Review of Mixed-Methods Research in Nursing. *Nursing Research,* 68**,** 464-472.

YU-CHIH, H., BACKMAN, S. & BACKMAN, K. 2010. Student Attitude Toward Virtual Learning in Second Life: A Flow Theory Approach. *Journal of Teaching in Travel & Tourism,* 10**,** 312-334.

YUNG, R. & KHOO-LATTIMORE, C. 2019. New realities: a systematic literature review on virtual reality and augmented reality in tourism research. *Current Issues in Tourism,,* 22**,** 2056-2081.

YUSOFF, R., ZAMAN, H. & AHMAD, A. 2011. Evaluation of user acceptance of mixed reality technology. *Australasian Journal of Educational Technology,* 27**,** 1369-1387.

ZACHARIADIS, M., SCOTT, S. & BARRETT, M. 2013. METHODOLOGICAL IMPLICATIONS OF CRITICAL REALISM FOR MIXED-METHODS RESEARCH. *MIS Quarterly,* 37**,** 855-879.

ZENG, J., XU, Y., SHI, Y. & JIANG, C. 2018. Inflammation role in sensory neuropathy in Chinese patients with diabetes/prediabetes. *Clinical Neurology and Neurosurgery,* 166**,** 136-140.

ZIEBER, M. & SEDGEWICK, M. 2018. Competence, confidence and knowledge retention in undergraduate nursing students — A mixed method study. *Nurse Education Today,* 62**,** 16-21.

ZULFIKAR, T. & MUJIBURRAHMAN 2018. Understanding own teaching: becoming reflective teachers through reflective journals. *Reflective Practice,* 19**,** 1.