A NEW REALITY IN TEACHING CHEMISTRY:
DEVELOPING A VIRTUAL REALITY INTERVENTION TO IMPROVE STUDENT UNDERSTANDING OF CONCEPTS OF CHEMICAL EQUILIBRIUM

by

Jason Ward

A dissertation

submitted in partial fulfillment

of the requirements for the degree of

Doctor of Education in Educational Technology

Boise State University

August 2023

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the dissertation submitted by

Jason Ward

Dissertation Title: A New Reality in Teaching Chemistry: Developing a Virtual Reality Intervention to Improve Student Understanding of Concepts of Chemical Equilibrium

Date of Final Oral Examination: 29 March 2023

The following individuals read and discussed the dissertation submitted by student Jason Ward, and they evaluated the student's presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Brett Shelton, Ph.D.	Chair, Supervisory Committee
Youngkyun Baek, Ph.D.	Member, Supervisory Committee
Yu-Chang Hsu, Ph.D.	Member, Supervisory Committee

The final reading approval of the dissertation was granted by Brett Shelton, Ph.D., Chair of the Supervisory Committee. The dissertation was approved by the Graduate College.

DEDICATION

I dedicate this to my wife, Moira, who has been by my side on this amazing journey.

ACKNOWLEDGMENTS

I am deeply grateful for the support and guidance of my professor and committee chair, Dr. Brett Shelton, and the members of my dissertation committee, Dr. Yu-Chang Hsu and Dr. Youngkyun Baek. Thank you for the time, effort, and wisdom that you have so generously imparted upon me throughout this process.

The faculty of the Educational Technology Department, who have taught me so much and worked with a sense of enthusiasm for this subject that was inspiring. Every course and every assignment during the past several years has led me to this point. I am in your debt.

Thank you to the members of my cohort, with whom I have struggled and persevered. I could not have asked for a more thoughtful, dedicated, and supportive set of classmates.

I would not have started, much less completed, this journey without the endless support and encouragement of my family. To my wife, Moira, and my children, Ewan and Nolan, words could never express my gratitude for everything you have done to help me reach this milestone.

ABSTRACT

The subject of chemistry is a cornerstone of high school science programs and a conceptual understanding of chemical equilibria and Le Chatelier's Principle continues to be both a fundamental and paradoxically difficult subject for educators and students. This study evaluated the effectiveness of a prototype immersive virtual reality environment. The prototype was designed and built to high school support students' understanding of the concepts of chemical equilibria and Le Chatelier's Principle. Data was collected to assess student understanding of the concepts being taught, as well as to evaluate the user experience with the prototype learning environment. Results of data analysis indicate that students were able to enhance their knowledge of chemical equilibrium and that the user experience was positive overall. The study contributes to the fields of educational technology, science education, and design of virtual reality learning environments. By collecting evidence to support the application of immersive virtual reality for this specific subject matter, it also provides a basis for future research utilizing similar environments for teaching and learning of other concepts as well.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEFINITION OF TERMS

Game (the game) – For the purposes of this project, "the game" is used to refer to the portion of the virtual reality application developed for this dissertation that includes gamified elements, such as challenges, scoring, and adaptive pathways based on user choices.

Immersive/Immersion - "a perception of being physically present in a non-physical world" (Freina & Ott, 2015, p.1)

Intelligent Tutoring System - a learning environment that adapts to the needs of individual students by applying algorithms and learning theories to maximize the development of new skills (Graesser, Hu, Nye, & Sottilare, 2016)

Le Chatelier's Principle - an empirical observation of the behavior of equilibria when they are disturbed (Lucanus, 2011)

Learning environment – For the purposes of this project, "learning environment" refers to the virtual world that exists in the application developed for this dissertation.

Prototype – For the purposes of this project, "prototype" refers the application developed for this dissertation work.

Virtual reality - any situation in which the individual experiences sensations other than those which are in their immediate environment (Blascovich & Bailenson, 2011)

Virtual world – a world created by a computer that simulates an environment in which some entities can be acted upon by the user (Bartle, 2004)

CHAPTER ONE

Introduction

There are many concepts in high school and undergraduate chemistry that are difficult for students to grasp, due to a combination of underlying misconceptions and the fact that many of these phenomena cannot be observed directly. However, chemistry is a fundamental science, critical to developing an understanding of the physical universe.

A key element of chemistry education is the concept of chemical equilibria and Le Chatelier's Principle. These topics are prominently described in the Next Generation Science Standards. However, these concepts are also among the most difficult for students to understand (Ghirardi et al., 2014). In fact, both textbooks and instructors have been shown to struggle in their own understanding of these concepts (Cheung, 2009).

Researchers and educators have attempted to resolve these issues for decades (Ghirardi et al., 2014). Approaches involving sequence of teaching, analogies, simulations, and inquiry-based learning have all been applied to the problem (Annisa & Rohaeti, 2018; Banerjee, 1995; Fogarty, Geelan, & Mukherjee, 2012; Ghirardi et al., 2014). It has been argued that attempts to teach Le Chatelier's Principle may even be detrimental to students' understanding of chemical equilibrium (Cheung, 2009; Lucanus, 2011).

There is a clear need for interventions that can improve students' understanding of chemical equilibrium. The concepts are both critically important and exceedingly complex. Research shows that misconceptions related to abstraction, terminology, and the fundamental understanding of chemical reactions all may play a role in these difficulties (Annisa & Rohaeti, 2018; Ghirardi et al., 2014). This study developed and tested an intervention to overcome these challenges, utilizing the emerging technologies of immersive virtual reality and intelligent tutoring systems.

Background of the Study

There is a large and growing body of research on the implementation of simulations, virtual reality, and immersive virtual worlds, and intelligent tutoring systems for teaching and learning (Freina & Ott, 2015; Hew & Cheung, 2008; Passig, 2015). Prior research has been conducted across a wide range of age ranges and content areas. However, the bulk of the research to date has been conducted in higher education settings (Bursztyn et al., 2015; Wang et al., 2015), using desktop virtual worlds and simulations (Basur & Durmus, 2010; Finkelstein et al., 2005; Trey & Khan, 2007). In addition, much of the current research on immersive virtual reality has focused on life science (Wilensky & Reisman, 2006), medicine (Butt, Kardong-Edgren, & Ellertson, 2018), and physics (Finkelstein et al., 2005).

While the body of literature on immersive virtual reality is growing, there remains to date a need for further research on immersive virtual reality for teaching and learning of science concepts at the secondary level. There have been no studies published to date that explore the use of immersive virtual reality to support the study of chemistry at the secondary education level. This study sought to add to the literature by investigating the use of immersive virtual reality to enhance students' understanding of a specific chemistry topic.

Research conducted to date has produced mixed results. However, there is evidence to support the effectiveness of virtual reality as a teaching and learning environment in specific circumstances (Dalgarno & Lee, 2010). For example, virtual reality can address issues of distance, cost, safety, and ethical considerations that would otherwise make learning activities unfeasible (Bursztyn, Walker, Shelton, & Pederson, 2017; Finkelstein et al., 2005; Mei & Sheng, 2011). Many concepts in chemistry present a combination of challenges that include safety, cost, and phenomena that are impossible to observe directly. Immersive virtual reality presents the opportunity to overcome all these obstacles, and this investigation will attempt to take advantage of those factors.

Purpose of the Study

The subject of chemistry is a cornerstone of science programs and curriculum standards across state and federal guidelines (Massachusetts Department of Elementary and Secondary Education, 2016; National Research Council, 2012; NGSS Lead States, 2013). A conceptual understanding of chemical equilibria and Le Chatelier's Principle continues to be both a fundamental and paradoxically difficult subject for educators and students. Many of these topics are difficult for students to grasp, for a variety of reasons, as discussed above.

This study evaluated the effectiveness of a prototype immersive virtual reality environment. The prototype was designed and built by the researcher to support students' understanding of the concepts of chemical equilibria and Le Chatelier's Principle. The environment takes advantage of the affordances of virtual reality, including interactivity and visualizations that are not possible without the use of technology. The prototype environment was created utilizing interactive simulations to represent chemical reaction systems. Students were able to interact with the systems, changing various aspects of the simulation, and were able to see the dynamics of chemical equilibria in real time. By testing students' understanding of these concepts before and after their experience with the virtual reality environment, quantitative data on student performance and user experience were collected and analyzed to assess the effectiveness of the environment as a teaching and learning tool.

The study contributes to the fields of educational technology, science education, and design of virtual reality learning environments. By collecting evidence to support the application of immersive virtual reality for this specific subject matter, it also provides a basis for future research utilizing similar environments for teaching and learning of other concepts as well. The driving factor behind the creation of the environment is to improve students' understanding of difficult chemistry concepts. It is hoped that this intervention will become a tool that educators can apply to improve learning outcomes for all students of chemistry at the secondary level. The study also examined the specific experiences of students with various elements of the simulated environment. This analysis will provide evidence for future immersive virtual reality environments across a wide range of subjects.

Research Questions

One of the challenges facing teachers and students concerning chemical equilibria is that the phenomena involve interactions that cannot be observed directly. While this problem can be alleviated by conducting experiments that produce visible changes (i.e., changing colors), such activities require the use of chemical reagents and equipment that require safety measures for use and disposal, which creates costs for the school. In addition, these reactions do not always proceed as described, due to student errors or other external factors. These errors can further confuse students or reinforce existing misconceptions.

Since one of the issues that generates problems for students is a lack of understanding of chemical reaction mechanics, analogies are often employed to help students visualize the changes that are occurring within the reaction system. Unfortunately, many of these analogies are incomplete or inaccurate, calling upon other science subjects or events from daily life. When the differences between the analogies and the actual concepts of chemical equilibrium are not explained clearly, further misconceptions and misunderstandings can be created.

Finally, educators may incorporate the use of technology and simulations. Such activities can create safe, consistent results with reactions that work exactly as described. The integration of technology can also enhance the experience by providing visualizations and graphical representations of reactions that are not possible with standard laboratory equipment.

This study incorporated elements of all three of the above practices to design, develop and test a prototype of a virtual reality learning environment. The study was driven by two main objectives. The first was the development and piloting of the prototype learning environment itself. The application was designed and created for this study by the researcher using commercially available development software and off-the-shelf virtual reality hardware. Research questions and data collection were designed to assess the performance of this prototype, both in terms of its ability to enhance a student's understanding of the chemistry concepts involved and the user experience with the learning environment itself. All of this data will be utilized in the development of future iterations of the learning environment.

The primary research question explored in this study was:

Research Question 1: To what extent do high school chemistry students learn concepts of chemical equilibrium via the implementation of an immersive virtual reality environment?

The data collected to address this question included pre- and post-test performance related to students' understanding of chemical equilibrium concepts. The details of the methodology and data collection will be provided in greater detail below.

In addition, data analysis was conducted to look for evidence that may relate to the following additional questions related to the design of the intervention itself:

Research Question 2: What technical and conceptual challenges do students encounter while using a virtual environment for learning chemical equilibrium concepts?

Research Question 3: What aspects of usability and user experience correlate to student performance?

Research Question 2 and Research Question 3 were addressed by a post-intervention survey, using a shortened version of the Virtual Reality Neuroscience Questionnaire (Kourtesis, Collina, Doumas, & MacPherson, 2019), along with adapted items from the System Usability Scale (Brooke, 1996) and the User Reaction Survey (Butt et al., 2018). A combination of items from these three instruments was administered to participants to create an overall user experience survey that was taken after completing the post-test assessment. These questionnaires use a Likert-type scale to assess user experience in various aspects of immersive virtual environments. The usability survey was analyzed to determine the overall usability of the prototype and satisfaction with the user experience. A correlation analysis was conducted to identify any specific user experience and usability items that may correlate to differences in student performance. The results of these analyses for Q2 and Q3 will be useful to guiding future iterations of the learning environment and may provide insights for other researchers and software developers. An overview of the conceptual framework for the project is presented in Table 1.

Table 1 **Overview of Conceptual Framework for Current Study**

Research Problem	Research Question	Data Collection	Data Analysis	Potential Outcomes
Concepts of chemical equilibrium are challenging for high school chemistry students.	To what extent do high school chemistry students learn concepts of chemical equilibrium via the implementation of an immersive virtual reality environment?	Pre- and Post-Intervention assessment	Paired t-test comparison of pre-test and post-test data to evaluate success of the learning intervention	Students will have significantly improved their understanding of the concepts of chemical equilibrium
There is limited research on the implementation of immersive virtual reality learning environments for teaching K-12 science concepts.	What technical and conceptual challenges do students encounter while using a virtual environment for learning chemical equilibrium concepts?	Virtual Reality Neurological Questionnaire, System Usability Survey, User Reaction Survey, Pre- and Post-Intervention assessment	Minimal and Parsimonious cutoff scores to determine suitability of the prototype. Minimum Scoring of System Usability Survey. Analysis of Mean Scores on User Reaction Survey Items for Positive and Negative Reactions.	User experience scores will guide future iterations of the prototype.
There is limited research on the implementation of immersive virtual reality learning environments for teaching K-12 science concepts.	What aspects of usability and user experience correlate to student performance?	Pre- and Post-Intervention assessment, Virtual Reality Neurological Questionnaire, System Usability Survey, User Reaction Survey	Correlation analysis to test for correlation between specific user experience scores and learning outcomes	Correlations between user experience and learning outcomes may highlight positive or negative relationships for further exploration.

Significance of the Study

As noted above, the study created and tested a prototype for an immersive virtual reality environment for learning concepts of chemical equilibria and Le Chatelier's Principle. The study was conducted in several sections of a chemistry course at a medium-sized public high school. Given the relatively small sample size and specificity of the topic under study, the results may not be broadly generalized across topics, settings, and content areas. However, there are several audiences that may benefit from the results and outcomes of the study.

First, the study contributes to the growing body of research on immersive virtual reality as it pertains to education, teaching, and learning. If the results support the use of the prototype as a positive intervention for the teaching and learning of chemistry concepts, then there may be a wide range of other topics in physical and life sciences that could be further explored using similar methods. Positive outcomes would also provide further evidence for the applicability of immersive virtual reality in science education and education more generally. The use of virtual reality to visualize otherwise invisible and abstract concepts would also lend itself to exploration of similar interventions, whether in virtual reality or in other formats.

Second, since the study involved the design and testing of a prototype, the results should be of value to developers and designers of other virtual reality environments for education. The success or failure of various design elements will contribute to developing best practices for future products in education and elsewhere.

Finally, the study will be of interest to researchers in science education, educational technology, as well as educators themselves. The results will contribute to the existing literature on the teaching and learning of the concepts of chemical equilibrium and Le Chatelier's Principle. There have been many attempts to improve student understanding of these complex and fundamental concepts in chemistry. The prototype will combine aspects of technology, learning theory, teaching practice, and hands-on activities to improve outcomes for teachers and students.

Rationale for Methodology

Educational Design Research (EDR) is a method for studying the development and practical applications of educational interventions (McKenney & Reeves, 2012). The process of evaluating an educational problem, creating a prototype, and testing the intervention under real-world conditions is ideally suited to EDR (Plomp, 2009). Reeves and Amiel (2008) suggested that educational design research is well-suited to studying the impact of educational technology and may in fact address many deficiencies of other methodologies. While there are several research objectives that can be examined using EDR, the proposed study fits this model as an interventionist approach (Van den Akker, Gravemeijer, McKenney, & Nieveen, 2006).

Educational design research uses an iterative approach to the research, development, implementation, and evaluation of an intervention. This study gathered data on a prototype immersive virtual reality learning environment. Data collected included a pre- and post-test to assess whether students gained a better understanding of chemical equilibrium, as well as a post-intervention survey to assess usability of the virtual environment.

Educational design research is frequently used in evaluation of educational interventions and programs, across multiple disciplines (McKenney & Reeves, 2012;

Turucz et al., 2021). More specifically to the proposed intervention, Turucz et al. (2021) applied EDR to a virtual reality environment for emergency training. The authors created a detailed protocol for using EDR, driven by a need to produce practical applications to practical problems. As noted by the authors, EDR represents an excellent opportunity for the study and application of educational innovation. Zikas et al. (2020) used a prototyping process and user experience surveys to gather data on a proposed scripting language for VR development. While this example was not explicitly an EDR study, the process of design, implementation, and evaluation followed a similar structure to EDR and was applied to a VR environment.

Assumptions of the Study

The study relied on several assumptions as they pertain to the science content to be learned, the technology, and the setting of the study. Many of these assumptions are common to all educational technology research (Willis, 2001), while others are more specific to the proposed project. The key assumptions of the proposed study and the justification for each are described below.

First, it is assumed that chemical equilibrium and Le Chatelier's Principle will continue to be core concepts in the science curriculum at secondary schools for the foreseeable future. This assumption is based on the continued inclusion of these topics in science standards at both state and federal levels. Specifically, the Next Generation Science Standards include understanding of these concepts as criteria for measuring students' understanding of the interactions of matter. Students will be able to "Refine the design of a chemical system by specifying a change in conditions that would produce increased amounts of products at equilibrium...Emphasis is on the application of Le

Chatelier's Principle and on refining designs of chemical reaction systems" (NGSS Lead States, 2013, HS Chemical Reactions).

It is assumed that technology will continue to advance and spread, and specifically that immersive virtual reality will become more accessible to educational institutions over time. This progress has supported technological advancements, and the emergence in recent years of new and less expensive virtual reality technology. For example, the hardware requirements for implementing interactive, immersive virtual reality would until recently have required an investment of two to three thousand dollars per student. However, the recent introduction of lower-cost, freestanding systems, such as the Meta Quest (Facebook Technologies, 2018, 2019) offer these features at a cost of less than five hundred dollars per student. Simpler systems, such as Google Cardboard, provide an immersive experience for even lower costs, if students have access to compatible smartphones. It is assumed that this trend will continue as new, more powerful, less expensive technology is introduced.

Teaching and learning are not achieved through technology alone (Willis, 2001). It is assumed that the prototype learning environment is based on established teaching practices and curriculum standards. The technology applied in this study has been selected and was developed due to the advantages it offers over other teaching tools (e.g., interactivity and visualization). However, these advantages merely enable educators to more effectively apply strategies that have been demonstrated to be effective in other contexts as well.

The proposed methodology, educational design research, relies on the specific assumption that the problem being studied is complex and does not have a clear, known solution (Kelley, 2009). This assumption will be discussed in greater detail in the following chapter, and the details of the methodology will be described in more detail in Chapter 3.

Finally, it is understood that all other aspects of the teaching and learning during the study will be identical, to the greatest extent possible. This assumption is required to establish support for the effectiveness of the technology for improving learning outcomes. However, since the teaching and learning of the content will take place in multiple class sessions, with up to five different groups of students, it is impossible to establish complete control of all variables. This assumption is both fundamental and problematic when conducting research in an actual educational setting (Kirkwood & Price, 2013)

Chapter 1 Summary and Organization of the Remainder of the Dissertation

The key issues to be addressed in this study included an attempt to provide an intervention that can improve student understanding of the concept of equilibrium, while taking advantage of the affordances of immersive virtual reality. These chemistry concepts are vital but problematic for both educators and students at the secondary and higher education levels (Ghirardi et al., 2014). Despite issues with cost, safety, and the abstract thinking required to fully grasp concepts that are invisible to the naked eye, these concepts are central to state and federal standards, and cannot be removed from the curriculum (NGSS Lead States, 2013). Therefore, this study created a prototype virtual reality learning environment that will give students a “hands-on” opportunity to interact with chemical reaction systems and to directly observe the changes that occur in a dynamic state.

Research on the implementation of immersive virtual reality has been conducted across multiple disciplines and in a variety of educational settings (Freina & Ott, 2015; Hew & Cheung, 2008; Passig, 2015). The present study was intended to add to the current literature in two critical ways. First, prior research has been conducted predominantly in higher education settings, using desktop virtual reality or virtual worlds and simulations. This study was specifically designed to address the needs of secondary students, using an immersive and interactive virtual environment. Second, whereas current research on immersive virtual reality has focused on life science, medicine, and physics, this study applied to specific concepts in the chemistry curriculum.

There is evidence to support the effectiveness of virtual reality as a teaching and learning environment in specific circumstances (Dalgarno & Lee, 2010). In secondary chemistry education, issues of cost and safety must be considered (Bursztyn et al., 2017; Finkelstein et al., 2005; Mei & Sheng, 2011). Further, the specific concepts addressed in this study involve phenomena that are impossible to observe directly. It is with immersive virtual reality that the present study may be able to overcome these obstacles to teaching and learning.

This study designed and tested an immersive virtual reality environment to support students' understanding of the concepts of chemical equilibria and Le Chatelier's Principle. The primary research question, as noted above, was whether interacting with an immersive virtual environment would improve students' understanding of these key concepts. To assess this research question, student performance on a pre- and post-test was analyzed. The second research question, as to the technical and conceptual challenges students encounter when using the intervention, was assessed in the form of a post-intervention survey. Results from this survey were used to identify the most successful and least successful elements of the intervention, to guide future iterations of the learning environment.

The study followed an educational design research model for the development, implementation, and evaluation of the proposed intervention, which will be described in more detail in Chapter 3. This design takes advantage of pre-existing groups such as students in the researchers' chemistry classes for the current school year, rather than randomly assigned control and experimental groups (Oncu & Cakir, 2011). The generalizability of this design is limited since it is specific to the population of the students in the schools under study and features a relatively small sample size.

Educational design research is a form of design-based research. It uses an iterative model for the development and application of interventions in a practical setting (McKenney & Reeves, 2012). The development, implementation, and evaluation of a prototype immersive virtual reality intervention were executed as a meso-cycle of EDR. Student learning was assessed by means of a pre- and post-test model. The effectiveness and satisfaction of the user experience was assessed using a post-intervention survey.

Organization of the Remainder of the Dissertation

This first chapter has introduced the topic of the study, background on the problem addressed, and an overview of the study and the research design. The remainder of this dissertation is composed of four chapters. In Chapter 2, a thorough review of the literature on the subject will be provided, along with linkage to prior work and the theoretical underpinnings of the study. In Chapter 3, the methodology and design of the prototype will be discussed and justified in more detail. Chapter 4 will present the results of the study. Chapter 5 will discuss the results, limitations, potential future directions of the Educational Design Research project, and contributions to the field. Finally, the Appendices will include a range of other information not provided elsewhere in the proposal, including an IRB application and data collection tools that were used for the study.

CHAPTER TWO: REVIEW OF THE LITERATURE

Introduction and Background

The following chapter includes a review of the literature in several key areas relevant to this study. First, in the following section, a description of the theoretical background for the use of immersive virtual reality in teaching and learning. The remainder of the literature focuses on the content and the application of technology in teaching and learning. The review will define the key concepts being investigated, as well as the opportunity to contribute to the field by addressing a pressing need through the application of emerging technologies. While the teaching and learning of chemical equilibrium has evolved over several decades, and its importance as a topic in the study of physical sciences is well established, it remains a difficult concept for educators and students alike.

In recent years, the application of technology to chemical equilibrium education has been studied (Finkelstein et al., 2005; Trey & Khan, 2007). However, these studies focused on desktop environments and simulations. With the emergence of virtual reality as an educational tool, there is a need for research into the potential benefits of this technology to address teaching and learning of chemistry at the secondary level.

The body of research involving virtual reality and virtual worlds in education has grown considerably in recent years. The introduction of new technologies, and the continuing decline in the cost of virtual reality hardware has opened the study of interventions across a range of subject areas. This study contributes to the body of literature by exploring the application of this technology to the concept of chemical equilibrium. This combination of technology and subject matter has not yet been examined and represents a significant opportunity for advancement of the field.

This review will describe the background of the challenges and prior interventions to address the learning of concepts involving chemical equilibrium and Le Chatelier's Principle. The body of research in this area includes efforts to improve teaching practices, educator training, and curriculum related to chemistry education. Examples of these prior studies include recommendations for new structure in the order of teaching chemistry (Ghirardi et al., 2014), the use of analogies and simulations (Trey & Khan, 2007), better teacher training (Cheung, 2009), and addressing misconceptions around basic terminology and concepts (Primastuti & Atun, 2018). These studies provide several options for improving teaching and learning of chemistry. However, they do not include the use of immersive virtual reality.

Subsequent discussion in this chapter will include definitions of the technologies to be applied, as well as examples of specific applications of virtual reality in a range of disciplines and settings. While prior studies have produced results that support the implementation of various virtual reality technologies for teaching and learning, there have yet been no studies conducted on the use of immersive virtual reality for improving student understanding of chemical equilibrium. Studies have, however, demonstrated that the use of simulations and interactive technologies may improve learning outcomes in chemistry (Trey & Khan, 2007). The present study proposes to expand on this body of literature by applying immersive virtual reality to enhance the effectiveness of interactive simulations.

The chapter will conclude with a summary of the literature, the theoretical background for the study, and the proposed methodology. This summary will lead to the detailed presentation of the study itself. This detailed description will include the design research model, the setting and participant population. The methodology section will also include details on the hardware, software, and design of the prototype learning environment.

Conceptual Framework

The theoretical foundation for this study is found in constructionism, which describes learning as an active process in which learners construct their own meaning of the content being presented (Papert & Harel, 1991). The theory has been applied to the teaching and learning of science through technology (Wilensky & Reisman, 2006). Dass, Dabbagh, and Clark (2011) found that studies on the use of virtual reality in education frequently include constructionism and constructivism as pedagogical foundations. Virtual reality provides learners with the opportunity to construct concepts in environments that allow for modeling, investigating, and exploring concepts that would otherwise be too abstract or impractical in traditional settings. In their work on a constructionist approach to computer programming, Noss and Clayson (2015) describe the benefits of layering abstract and scientific principles, which students can investigate at a variety of levels. The current study followed a similar approach, by giving learners the opportunity to manipulate a range of factors to construct their own understanding of chemical equilibria and Le Chatelier's Principle.

Teaching and Learning Concepts of Chemical Equilibrium

The following section presents a review of the literature in several areas related to the current study. First, an exploration of the history, challenges, and prior research into the teaching of chemical equilibrium across a range of settings and teaching strategies (Davenport et al., 2014). This section highlights the need for further research and additional interventions for the teaching and learning of this fundamental concept in chemistry. This section includes examples of the use of educational technology for teaching equilibrium, including simulations. However, prior research has not specifically applied immersive virtual reality to this topic.

The remainder of the literature review discusses prior work in the areas of virtual reality and intelligent tutoring systems for teaching and learning. There is a large and growing body of research in these areas, and evidence to support the use of immersive virtual reality as a teaching and learning tool. Previous studies have applied various virtual reality and simulation technologies for the teaching of science concepts. It is proposed that these studies provide evidence for the possibility that immersive virtual reality may offer advantages for learning concepts of chemical equilibrium.

There is a long history of research into the challenges and potential solutions for the teaching and learning of chemical equilibrium (Banerjee, 1995). There have been several areas of interest identified and tested that give rise to these difficulties. Students and teachers have pre-existing misconceptions about what takes place in a system at equilibrium (Annisa & Rohaeti, 2018). There are multiple, and at times contradictory approaches to answering questions or making predictions about equilibrium (Lucanus, 2011). The topic itself requires a knowledge of a range of other chemistry concepts. The terminology involved is sometimes misinterpreted or misused by teachers and students alike. Mathematical interpretations that may assist in understanding require skills that are beyond the level of some secondary and tertiary students, or with which they are not proficient (Cheung, 2009). Finally, it is impossible to observe the phenomena directly. Even when models are used, or macroscopic features of changes (i.e., color) can be utilized, these examples offer, at best, an incomplete picture of what is taking place at the atomic and molecular levels (Maia & Justi, 2009).

Tyson and Treagust (1999) highlighted a number of these issues in their study. Utilizing a combination of two-tiered testing, observations, and interviews, the authors found that there were several areas of difficulty for secondary and tertiary students in understanding equilibrium. Further, the testing showed that, even when students were able to correctly predict the outcome of a scenario, they often failed to correctly identify the underlying cause of that outcome. This disconnect in understanding is yet another problem with teaching and learning of these concepts. Educators frequently use examples and models to build understanding and make connections to daily events. However, these examples and models have limitations and imperfections that may reinforce misconceptions, even when the limitations are clearly delineated by instructors.

Primastuti and Atun (2018) note that one of the deficits in current science education worldwide is a focus on the theoretical, with a lack of practical, hands-on learning to reinforce scientific concepts. Chemical equilibrium, as previously noted, presents a particular challenge in this area, since the interactions taking place at the molecular level are impossible to observe directly. The authors examined *science literacy*, defined here as “the intersection between what people know and what they can do," (p.1) as an indicator of success in understanding chemical equilibrium. Two of the three key domains of science literacy described in the study include content knowledge and the ability to explain scientific concepts.

The study followed a descriptive quantitative design (Primastuti & Atun, 2018), and involved a series of questions in several domains of science literacy. The results indicated that all the students in the study (Grade 11 secondary school students) demonstrated low science literacy scores. Of particular interest to the proposed study, questions relating to explanation of equilibrium, interpreting the data, and content knowledge all had scores below 35%. While solutions or recommendations for improving these scores were beyond the scope of the study, the results clearly support the need for interventions that could improve student success in these areas, as they relate to chemical equilibrium.

The issues related to teaching and learning of chemical equilibrium extend beyond the students, proving difficult for many educators to address as well. Chani, Ngcoza, Chikunda, and Sewry (2018) observed and interviewed highly qualified teachers of high-achieving students and found that instructors face many of the same challenges as their students. Even though the instructors may possess the content knowledge of the subject, the complexity and abstract nature of the topic make teaching difficult. For example, while students may have prior knowledge of reactions and equilibrium, the level of understanding and misconceptions make it difficult to draw upon that knowledge. In fact, one of the interventions that instructors frequently apply (i.e., analogies) may only serve to introduce new misconceptions and confusion among students who have preconceived analogies or models. The study demonstrated that these issues are similar for high-achieving students and well-qualified teachers.

Another study involving teachers' understanding of equilibrium concepts was conducted by Cheung (2009). The focus of the study was on the limitations, and potential negative impact, of Le Chatelier's principle on both teachers' and students' understanding of equilibrium. In a quantitative examination of teachers' answers to equilibrium problems, as well as a qualitative analysis of their explanations for those answers, Cheung found that nearly 90% of teachers surveyed failed to solve an equilibrium problem correctly. Many of the teachers used variations on Le Chatelier's Principle to justify their incorrect answers. Beyond merely demonstrating the difficulty of the content, Cheung proposed that Le Chatelier's Principle is incomplete, misunderstood, and perhaps damaging to students' ability to understand concepts of chemical equilibrium.

In a similar article, Lucanus (2011) proposed that the complexities and limitations of Le Chatelier's Principle are such that it may best serve students to teach equilibrium by other methods. While the proposal does not include any experimental evidence, the author suggests that equilibrium would be better understood if students are given a better understanding of the concept in terms of kinetics and particle interactions. While these recommendations may provide a more logical and concrete explanation for the behavior of systems at equilibrium, it does not address the fact that these interactions take place on a scale that makes such explanations abstract.

Some researchers have advocated for changing the sequence of learning concepts, to overcome the challenges of chemical equilibrium. Fogarty, Geelan, and Mukherjee (2012) conducted a study to determine if understanding could be improved by using visualizations in advance of receiving the traditional teaching methods concerning chemical equilibrium. The authors note that, regardless of teaching sequence, the concepts of equilibrium and Le Chatelier's Principle must be taught carefully to avoid misconceptions and incorrect predictions. Participants were divided into two groups, with one group accessing an online visualization tool individually and independently, followed by more traditional, teacher-led teaching and discussion. The second group received the teacher-led discussion first, followed by independent work with the visualization.

The study used a multi-layered pre- and post-test model for quantitative analysis. All participants were given a brief pre-test prior to any teaching or using the visualization. Students in both groups were then tested again following their first session (either visualization or teacher-led), and then tested a third time following their second session. The results showed that, by many measures, the sequencing did not impact students' learning. Student improvement also showed no significant difference between achievement following the visualization or the teacher-led discussion. In other words, both methods appeared to have equal impact on student learning.

This study had several limitations, both in general and as pertains to the current study. First, as noted by the authors, the sample size was small, and some subgroups were too small to generate valid statistical analysis (Fogarty et al., 2012). Further, the visualization used in this study had limited interactivity, as described by the authors. Students were able to observe multiple representations of content and could pause and rewind the visualizations. However, the content was primarily presentational, rather than interactive. Second, while the visualization was computer-based, it was not immersive.

Several studies have been conducted involving modeling, inquiry-based learning, and educational technology tools to improve outcomes for teaching and learning of chemical equilibrium. concerning different approaches to teaching as noted above, Fogarty et al. (2012) studied a computer-based visualization tool and found it to be equally successful as teacher-led discussion in building conceptual understanding. The studies described below explore methods for creating practices and sequences based on a constructionist approach to dealing with the complexity of chemical equilibrium.

Ghirardi et al. (2014) developed a sequence of activities to assist students in developing an understanding of equilibrium by first representing the equation symbolically, and then carrying out the experiments. The experiments exhibited a variety of characteristics that facilitated discussion and stepwise understanding of equilibrium and incomplete reactions. This sequencing and modeling approach is similar to the current project, in that it allows the learner to construct their knowledge through multiple steps of varied types. It also provided multiple modes of accessing the information, including representing the reactions symbolically (chemical equations), visually (diagrams), and practically (hands-on lab work). By building knowledge iteratively in this way, and holding class discussions following each activity, the authors suggest that students could work through the logic and underlying factors that drive chemical equilibrium problems. While the observations made by the authors supported their approach, the project did not follow an experimental model, and did not objectively measure performance.

Maia and Justi (2009) highlight the importance of having students construct models of events, especially in chemistry. As the authors note, because many of the key concepts in chemistry "are essentially abstract, modeling and the use of models enables chemists to visualize entities and processes, to think about questions related to them" (p. 603-604). The authors followed a similar model to that of Ghirardi et al. (2014), by allowing the students to work through a series of problems that incrementally introduced new concepts related to equilibrium. Maia and Justi went further, in that their modeling strategy required students to make predictions based on prior models, test those predictions, and revise their models and understanding to explain new observations. Students' models were created and revised using a combination of clay, ball-and-stick representations, drawings, and discussion. Due to the limitations of prior knowledge of the students in the study, the predictions and concepts being taught included only the qualitative aspects of chemical equilibrium. This study deliberately did not use technology, but the concept of building and revising a model based on a series of increasingly complex reactions supports the concepts planned for the current study.

In a quasi-experimental study conducted by Annisa and Rohaeti (2018), students followed an inquiry-based model in developing their understanding of chemical equilibrium. The activities were guided, but allowed the students to formulate hypotheses, test their understanding, and analyze data to determine whether their model was supported. The results of a pre- and post-test analysis showed that students in the treatment group (who followed the inquiry-based model) showed significantly more growth in their understanding of chemical equilibrium, compared to the control group.

The design of this study, compared to the qualitative results from Maia and Justi (2009) and Ghirardi et al. (2014), provides quantitative data to support the use of this constructionist, iterative approach to learning concepts of chemical equilibrium. However, this study did not apply technology, and relied on observations at the macroscopic level (using hands-on experiments) to develop understanding. As previously noted, these macroscopic observations aid in making reactions less abstract, but they are limited in their power to provide a detailed understanding of what is taking place at the particle level of chemical reactions.

Akkus, Kadayifci, Atasoy, and Geban (2003) carried out a similar, quasi-experimental study based on constructivist principles for teaching and learning of chemical equilibrium concepts. Unlike Annisa and Rohaeti (2018), however, the work of Akkus et al. did not involve actual hands-on laboratory experiments. Rather the treatment group made predictions, compared results to those predictions, identified misconceptions and errors, and iteratively changed their model of equilibrium until an understanding was achieved. The authors found that students in the treatment group performed significantly better than those in the control group in a pre- and post-test comparison.

Finally, Trey and Khan (2007) studied the application of computer-based analogies to improve understanding of chemical equilibrium. The study focused on the opportunity presented by computer-based models to represent Le Chatelier's Principle dynamically and visually by presenting students with the analogy of a set of scales. This imagery is a common analogy used in the teaching and learning of dynamic equilibrium. The study followed a quasi-experimental, post-test only design to examine the impact of the computer-based analogy.

The model used in this analogy is like the current study in two key areas. First, it uses analogies and dynamic representations of chemical reactions to aid in understanding. Second, students could interact with the simulation, changing various factors in the system, and observing the corresponding change in the analogy, along with graphical representations of concentrations within the system. The control group in the study used a computer simulation that did not include the analogy model. The results of the post-test analysis showed a significant improvement in achievement for the treatment group over the control group. This data supports the use of interactive analogies for improving understanding of chemical equilibrium concepts. The authors suggest that future research would be beneficial in understanding the impact of dynamic computer-based analogies, which the current study was designed to address.

The Horizon Report (Freeman et al., 2017) includes authentic learning experiences as one of the key trends and challenges in the evolution of teaching and learning. Authentic learning refers to the application of real-world skills in real-world situations, to enhance the acquisition and transfer of new knowledge. One of the factors that the Horizon Report points to is the lack of real-world environments in current educational settings. While partnerships and field trips to academic or industrial settings offer the opportunity for students to have authentic learning experiences, there will always be limits, in terms of both logistics and liability (Bursztyn et al., 2017). Virtual environments overcome these obstacles by allowing educational institutions to build immersive, artificial constructs to simulate nearly any activity required. These activities can then be carried out, adapted, and personalized to the needs of individual students.

Teaching and Learning with Immersive VR and ITS

There are many definitions and many different models of virtual reality (VR). The Oxford English Dictionary defines VR as “The computer-generated simulation of a three-dimensional image or environment that can be interacted with in a seemingly real or physical way by a person using special electronic equipment, such as a helmet with a screen inside or gloves fitted with sensors.” (“Virtual Reality”, 2018). The Horizon Report (Freeman et al., 2017) includes VR as one of the key trends emerging in education and describes the technology as simulating sensory experience of physical presence and interaction with people or objects in an environment generated by a computer. Bursztyn et al. (2015) described this perception as the sensation of “being there” (p. 93) in their description of an augmented reality field trip, which can also be applied to the sense of presence in a virtual environment. Blascovich and Bailenson (2011) argue that VR requires no technology whatsoever and can be broadly applied to any situation in which the individual experiences sensations other than those which are in their immediate environment (i.e., daydreaming). The authors propose that “the brain often fails to differentiate between virtual experiences and real ones” (p.1). One of the benefits of immersive virtual reality is that it allows for students to investigate and have authentic experiences in a safe environment that can be customized to the needs of the individual student (Kirkley & Kirkley, 2007).

The concept of immersion is likewise broadly defined and refers to the level at which the user is aware of the real versus the virtual. This awareness is described by Bossard, Kermarrec, Buche, and Tisseau (2008) as providing the user with the sense that they exist within and can influence the virtual environment. Freina and Ott (2015) provide a very simple definition of immersion as “a perception of being physically present in a non-physical world” (p. 1). There are many factors that influence the level of immersion, from visual to auditory and tactile sensory perception. The greater the level of immersion, the more likely the user is to perceive the environment as reality.

Technology to allow students to experience immersive virtual reality is becoming more accessible and will continue to expand in both the consumer and educational technology field in the coming years, as prices continue to drop and adoption in schools drive the creation of new content (Kinshuk, 2016; Nelson & Erlandson, 2012). There is a need for more research on the educational benefits of immersive virtual reality. However, there is evidence that virtual reality has both significant and lasting effects on learning (Ahn, Fox, Dale, & Avant, 2015; Passig, 2015).

The key components of VR, for the purposes of this project are (a) simulating the sensory experience of a physical environment, (b) allowing users to interact with objects in the environment, and (c) a sense of immersion, in which the user does not distinguish between the perceived virtual environment and reality. While the project does not require the highest levels of immersion to succeed as a proof of concept, every effort will be made to enhance the user experience in terms of increased immersion. This vision for the system is a factor in choosing the hardware and software components of the project, as described below.

While virtual worlds have existed for some time, in the form of desktop worlds (e.g., Second Life), there is reason to believe that immersive virtual worlds may have additional benefits. Ahn et al. (2015) tested the effects of immersive virtual reality versus a computer-based virtual world in the immediate and longer-term effects of a message on environmental consciousness. Results showed that, both immediately after delivery of the message and 1 week after the message delivery, the high interactive group experienced greater behavioral modification and environmental awareness.

The term *intelligent tutoring system* refers to a learning environment that adapts to the needs of individual students by applying algorithms and learning theories to maximize the development of new skills (Graesser, Hu, Nye, & Sottilare, 2016; Perez, Skineer, & Chatelier, 2016). Simple versions of ITS guide learners to one of several different learning paths, based on pre-test performance and formative assessments. These systems apply different questioning strategies and offer student choice. However, an intelligent tutoring system can play an active role in guiding students, both in selecting appropriate content, and in providing formative feedback (Gonzales-Calero, Arnau, Puig, & Arevalillo-Herraez, 2015). Such systems will develop and dynamically adjust learner profiles that impact the future behavior of the system.

A more sophisticated intelligent tutoring system was presented by Schiaffino, Garcia, and Amandi (2008). The ITS gathers data about a student by observing the student's interactions and performance during an online course. Activities such as performance on exams, time spent on studying certain topics, and the number of practice problems completed. The ITS uses Bayesian networks to develop a profile of the student's learning style. Once a profile is built and stored by the ITS, the system then can actively engage with the student, when it detects a situation that might be challenging, and the ITS can make suggestions or prompts to guide the student to the most appropriate course of action. This type of ITS can also uses machine learning to analyze aggregate student performance and further refine future predictions and adaptations.

An ITS that builds student profiles can provide feedback dynamically, based on parameters set by the instructor, combined with the context of the learning situation, and the student's profile. If, for example, the ITS determines that a student should read the material about a topic in a certain order, and the student is reading out of order, the ITS recommends going back to the relevant prior content. The student also could accept or reject the recommendation of the ITS. In either case, the ITS stores the student's feedback and applies this information when making recommendations in future situations that are similar. Results of survey and usage data in the sample study were positive, based on both the activity logs and student feedback. Most recommendations made by the ITS were received favorably.

At their core, virtual worlds created for both education and entertainment have a fundamental element based on learning. Whether the purpose is to teach specific content, social skills, specific game-based tasks, or simply exploration of identity, virtual worlds are collaborative, motivational, and generally intuitive to use (Bartle, 2004).

Van Eck (2007) describes the nine events required for teaching and learning of new information first proposed by Gagne (1965). The events include gaining attention, informing of the objective, recalling prior knowledge, presentation of instruction, offering guidance, opportunities for practicing new skills, feedback, assessment, and enhancement of retention and transfer. The focus of Van Eck's discussion was on digital game-based learning, but the same principles can be easily applied in the design of a virtual world integrated with an intelligent tutoring system (ITS).

Passig (2015) examined the benefits of immersive virtual reality on cognitive skill development, conducting a series of experiments to establish whether the acceleration of learning and cognitive skill development could be caused directly by interaction with an immersive VR environment. The experiments tested a variety of learning needs and showed that immersive virtual reality had the ability to produce significant improvements in understanding the experiences of others, improving inductive and deductive reasoning, and for aiding students with specific learning challenges. In each case, the immersive experience was superior to alternatives, including 2-dimensional virtual worlds.

The cognitive process of problem-solving is also at the core of simulations and immersive experiences (Baker & Delacruz, 2016). Students can analyze, plan, test, and revise strategies for solving the problems presented in the virtual world. The world can provide immediate feedback, and the learner can apply that feedback to the next attempt or the next task. This feedback allows students to engage with content in a practical, hands-on manner, rather than dealing with abstract concepts. This process is particularly valuable in the context of high school chemistry class, where many concepts are invisible to the naked eye, and in some cases are entirely abstract models based on observation. Immersion in a virtual world not only gives learners the opportunity to experiment in ways that would be impossible or impractical in the real world, but there is also the ability of the simulated world to demonstrate the abstract in a concrete manner (i.e., allowing students to step inside a model of an atom or molecule during a reaction).

For students to demonstrate their mastery of content, some form of assessment is required. The assessments may be informal, formative, formal, and summative. However, in the traditional educational model, many of these assessments require students to apply information in hypothetical scenarios, or by making calculations based on word-based problems. As noted by Rus and Stefanescu (2016), the traditional model of implementing a pre-test to assess prior knowledge, and a post-test to assess learning growth can as much as double the time it takes to deliver a lesson. Put another way, pre-tests and post-tests can easily account for more than half of the time spent in a learning activity, dramatically reducing the time a student must spend on learning new content.

In a virtual world, assessments are ingrained in the simulation itself. The environment provides opportunities for learners to demonstrate their knowledge by applying it to authentic tasks and problems. The environment and the ITS can provide feedback and guidance, based on the needs and results of each student's work, to maximize learning and transfer. Another benefit to having learners interact with the content in a virtual environment is that longitudinal data of every student behavior can be captured and analyzed for both immediate and future use. The intelligent tutoring system can assess a student's prior knowledge and cognitive abilities during the learning activity. This subtle form of assessment not only reduces the time spent on formal assessments, but also reduces the cognitive load on the student (Lintean, Rus, & Azevedo, 2008). Using computer-based environments also makes it possible to assess via non-testing methods, such as measuring time spent on specific tasks and the number of hints utilized by students during a particular task, that can provide an accurate measure of a student's ability (Rus & Stefanescu, 2016). This information can be used by an intelligent tutoring system to generate more personalized content and learning pathways for each student. This unobtrusive model of assessment is called *stealth assessment* (Shute & Ventura, 2009)

Constructivist theory suggests that the one-on-one dialogue and interaction between students and tutors are critical in the learning process. Artificial intelligence and machine learning can enable the development of intelligent tutoring systems that use natural dialogue, to create an experience like that of working with a human tutor or instructor (Rus, D'Mello, Hu, & Graesser, 2013). Natural language understanding will require an awareness of contextual meaning, but if combined with the creation of avatars within the simulation, this mode of communication will allow for more meaningful interaction between students and intelligent tutoring systems (Fakinlede, Kumar, Wen, & Kinshuk, 2013). An online example of this type of agent, called DeepTutor, was developed by Rus et al. (2013) as an intelligent tutoring system in the domain of physics.

Wang et al. (2015) developed an intelligent tutoring system to examine whether a personalized intelligent tutoring system (ITS) would be effective with students using iTutor vs. students who had access to the same material without iTutor. Participants were freshman college students in an information technology class. Students from four different classes were assigned to either an experimental or control group. Both groups received direct instruction, practice time, and had access to the same learning materials online. However, the experimental group was able to use iTutor along with the additional material, while the control group simply had access to the files.

Data showed that the experimental group performed significantly better on the post-test than the control group. Researchers suggest that the dynamic nature of iTutor allowed students to proceed at their own pace and on their own time. iTutor helped to offer the type of guidance that a human instructor may have done under different circumstances. Authors specifically note the benefits of ITS for distance and self-directed learning, and the benefit to teachers, by reducing some of the time required by individualized attention (Wang et al., 2015).

Bossard and colleagues (2008) highlight the potential of virtual reality to create experiences that enhance the transfer of acquired knowledge to new situations in the real world. Examples include the ability of medical personnel or emergency responders to experience crisis situations, and to develop the skills required for appropriate response, without the need to wait for or create these scenarios in the real world. Such events can be easily tailored, executed, analyzed, and reset as needed to achieve the desired outcome.

The current project involved the creation and testing of an immersive virtual reality learning environment. More detailed definitions of these terms, and of the system itself (e.g., structure, hardware, software) are provided below. However, the overall vision for the project was to create a user experience in which the student actively engages with a simulated environment, complete with auditory, visual, and tactile sensory input. The students demonstrate mastery of content and key skills in a series of learning activities that test their knowledge and abilities. The environment provided both remedial and advanced tasks to meet the needs of individual students, while ensuring that all students progress toward mastery of the essential content. Interactions with the system were both direct, in terms of feedback and assistance, as well as indirect, when the system alters the pathways followed by the student throughout the process in real time.

Connection between VR and ITS

Zhu, Yu, and Riezebos (2016) presented a research framework for smart learning environments. The authors note that the most effective environments place students in the central role of their own learning. In an immersive environment, the student is placed in a learning situation where they are the central actor in the environment. Combined with an ITS, the environment itself adapts to the needs of the learner, creating a personalized experience that meets the needs of each individual, which "provides the necessary learning guidance, suggestions, or supportive tools...in the right form, at the right time, in the right place" (p. 4). The authors point out that an ideal learning environment is one that provides an awareness of context, with adaptive content, timely feedback, and interactivity. All these features would be present in an immersive virtual environment with ITS.

Personalized learning environments also identify where the student is in their learning process, what support they need, and how they are most likely to receive that support effectively in real time (Kinshuk, 2016). Such environments adapt to match the ability of the learner, whether novice or advanced, by limiting or expanding the options available, or prescribing more or less intensive scaffolding for each task (Van Lehn, 2011). The VR and ITS environment can present a wide range of situations, targeted to the skills of each student, to enhance transfer of newly gained knowledge to different scenarios. The environment itself can be modified in real time to present new challenges, but only within the parameters met by each learner in previous learning experiences.

Another key concept that would be made possible or enhanced by the VR and ITS environment is the psychological experience of "flow." Flow is a state in which the user experiences "an exhilarating sense of control and mastery that can arise from pursuing a focused, goal-driven activity" (Bartle, 2004, p. 157). This experience is familiar to athletes and video game players, or anyone who has achieved this level of engagement with a particular subject matter.

A virtual reality environment with ITS constantly monitors the learner's level of ability, and to present challenges that stretch each individual's knowledge, while setting achievable learning goals. If properly implemented, the user is unaware of the support or modifications made to achieve these successes, which encourages further exploration of new content and skills. In the prototype system developed for this study, these support structures will take the form of visual representations, text explanations of outcomes, audio feedback in natural language, or completely behind the scenes, so that the learner is not even aware of the changes taking place. For example, regarding the learning of chemical concepts, such as stoichiometry, the system could present students with more or less complex reactions, so that the underlying concepts are present, but the reactions themselves may be different. This variety is like offering students of different reading ability books on the same topic, but with different levels of difficulty. However, in a chemistry lab, such personalization would be time consuming, expensive, and impractical. The students would not realize that their experience is different from another student. Each task or skill is introduced and practiced until it is mastered, and the next, more challenging task can be introduced. Each learner progresses at a rate that meets their own abilities, in much the same way that video games offer training or tutorials to new and experienced users (Becker, 2013).

Certainly, many of these concepts are not new. In fact, they represent best practices that have been known in education for years or decades. However, one aspect of the current system that is not feasible or even possible in the traditional setting is the dynamic adaptation of the environment itself. In a virtual environment the integration of the virtual environment with the intelligent tutoring system means that the environment itself can be dynamically updated to meet the needs of each individual student.

The ITS can prompt the virtual environment to generate an entirely different setting or sequence of activities, either to lower or raise the level of challenge. If a student needs specialized instruction, the environment can be altered to include additional elements that can assist with gaps in prior knowledge or make changes to satisfy accommodations for students with learning disabilities (i.e., text-based information, breaking down sequences into smaller chunks of content, more frequent checks for understanding). This combination of personalization and situational learning could have powerful benefits for students of all abilities and skills levels. The prototype environment will include a rudimentary ITS, including feedback, remedial scaffolding, and a branched system of presenting players with a variety of pathways depending on their success in each challenge. Future iterations will be developed to include sophisticated systems for evaluating and guiding learners along different learning pathways within the environment.

The emerging field of learning analytics has the power to dramatically enhance the learning experiences of individual students. Learning analytics involves “the process of collecting and studying usage data in order to make instructional decisions that will support student success” (Becker, 2013, p. 63). With the ability of virtual environments to collect and track data on a wide range of student behaviors, the actions of learners can be utilized to create both enduring and dynamic profiles (Kinshuk, 2016). These profiles, in turn, can be applied to intelligent tutoring systems. On a larger scale, as more data becomes available, advanced data mining can be utilized to find patterns in student behavior that can be applied to individualized learning paths (Johnson et al., 2014). Other metrics, such as number of attempts, as well as more traditional pre- and post-test or formative assessment models can also be used to measure cognitive skill and to establish a baseline for both cognitive ability and prior content knowledge (Chen, 2009; Hwang, 2003). Diagnostic data can be collected in a more natural manner, via student/ITS interactions using avatars, in-world support dialogue, and changes to the environment itself (Rus & Stefanescu, 2016).

Virtual Worlds and Learning

In their review of the literature, Hew and Cheung (2010) described three key areas in which virtual worlds were studied for educational purposes: social interaction or communication, creating a sense of space, and allowing users to interact within a virtual space. While the studies examined for the literature review described “immersive” virtual worlds, it is important to distinguish between levels of immersion, for the purposes of this discussion. Immersion refers to the degree to which a user feels that they are in an environment (Bartle, 2004). This sense of presence can be achieved in a variety of ways, using visual, auditory, and tactile cues. In many of the studies described here, immersion is achieved to varying degrees by creating rich, three-dimensional representations of physical environments. The sense of user presence, in some cases, is enhanced by the introduction of an avatar - a visual character used to represent the user within the 3D environment. The user controls the movement, appearance, and actions of their avatar, giving them a presence by proxy within the world.

Examples of these types of environments include Second Life and Active Worlds, as well as commercial video games, such as World of Warcraft and Minecraft. These virtual worlds share several common features, including communication systems, the ability to create and modify or act upon objects, and some degree of freedom of choice in terms of where to go and what to do. It is these features that have drawn the interest of researchers looking to explore the elements of teaching and learning (Brom, Pruess, & Klement, 2011; Callaghan, et al., 2009).

Another definition of immersion involves the ability of users to act upon and manipulate simulated environments, such as circuits or machinery, for the purposes of training or education (i.e., Physics Education Technology). These environments typically exist as 2D or 3D simulations on a computer screen, and they provide users with the opportunity to explore or learn about the workings of specific devices, or to train for emergency situations (Basur & Durmus, 2010). These environments have the benefit of being able to experience failure conditions without compromising the safety of the user or the real-world environment. They can also include other features that allow for educational support within the environment or can display information that would not be visible in the real world (i.e., the flow of electrons in a circuit).

In other examples of immersion, users experience a first-person sense of being in the environment. This level of immersion can be achieved by some combination of a head-mounted display (HMD), audio equipment, motion tracking, and tactile feedback. Compared to virtual worlds, such as Second Life, in which the user sees a representation of themselves within the space, this immersion technology gives the user a sense that they are physically present within an environment. This sense of presence, along with the ability to move around and interact with objects in the virtual environment, can generate a situation in which the user is cognitively unable to distinguish between reality and the virtual world (Blascovich & Bailenson, 2011).

Stefan (2012) further describes these immersive environments, in which the user can interact with an environment where the normal rules of reality may or may not apply. These features provide developers, teachers, and students with opportunities to explore concepts that would otherwise be impossible or impractical. They can also offer freedom of choice, which encourages the learner to take ownership of their experience. There are tools for communication and collaboration, or for individual exploration. The level to which these environments create a sense of presence is determined by several factors, including the technology through which the user interacts with the world. A key feature highlighted by Stefan is the ability of virtual environments to create experiential learning opportunities for students (Rubio-Tamayo, Barrio, & Garcia, 2017).

Bombari, Mast, Canadas, and Bachman (2015) examined the application of "virtual humans" or representations of humans in immersive virtual environments. The authors highlight several uses for these virtual tools in education, training, and social constructs. Among the aspects discussed are the ability of virtual humans to be controlled by specific constraints within the context of the learning environment. An example of this flexibility is that the virtual human can be tailored to the specific needs of an individual learner.

For example, Blascovich and Bailenson (2011) describe studies in which learners are more open to communicating with virtual humans who are subtly designed to appear similar to the learner. Bombari et al. also cite an example of this type of interaction in which women are more engaged when representations of female role models are present in the space. One key example of virtual human use cited by Bombari et al. (2015) is the use of virtual humans to provide feedback and guidance. In the current study, a virtual human would be the manifestation of the intelligent tutoring system in the virtual space, which could be a feature of future iterations of the prototype.

This discussion will examine studies that fall into one of four loose categories: studies using open virtual worlds (Second Life, Active Worlds), studies using commercial games as virtual worlds (Minecraft, World of Warcraft), studies using simulations and virtual environments created specifically for teaching and learning (Physics Education Technology), and studies using more immersive technologies, such as HMDs, motion tracking, and tactile feedback. Each study will be reviewed, the research methods will be discussed, and results of the studies will be examined.

The two most frequently referenced open virtual worlds are Second Life and Active Worlds. These worlds are three-dimensional environments in which the user is represented by an avatar. The user can control the actions and appearance of the avatar, and can create, manipulate, and interact with objects in the environment. These worlds also incorporate communication tools, such as text-based chat, voice chat, and nonverbal communications (i.e., avatar gestures).

Callaghan et al. (2009) explored the educational possibilities afforded by Second Life, an online, three-dimensional virtual world, in which users can interact with others, as well as interact with the world itself. One example of the educational application of virtual worlds discussed by the authors is the creation of Engineering Island, an area within Second Life created and maintained by the University of Ulster. This environment allows students and teachers to meet in the virtual space online, regardless of their physical proximity in the real world. In the virtual environment, users can communicate, collaborate, and interact in a virtual laboratory, where they can observe several demonstrations of engineering concepts.

In the case of the University of Ulster virtual space, one of the key features explored by Callaghan et al. (2009) is the integration of Second Life with a learning management system. As the authors note, one of the drawbacks to using an open virtual world for education is that, by definition, the open world does not have explicit learning objectives or prescribed actions. While this potential challenge can be mitigated by the presence of and guidance of an instructor in the space, it does not directly support the collection of student performance data or assessment. The authors describe the use of a tool called Sloodle, which allows for the connection of Second Life to a learning management system, which can then collect and store student data automatically.

The work of Callaghan et al. (2009) was exploratory in nature, and its descriptions offer more of a proof of concept than any formal research design. However, the conclusions drawn by the authors suggested that the integration of an open virtual world, such as Second Life with a learning management system could provide teachers and students with the ability to explore concepts and collaborate in ways that would otherwise be impractical or impossible.

Weir (2008) used anecdotal evidence to demonstrate the power of virtual worlds to engage students in inquiry-based activities and situate learning experiences in virtual versions of real-world settings that the students would otherwise be unable to visit. The author describes a specialized version of Second Life (Teen Second Life) that educators use to enable students to explore issues of citizenship, sustainability, and responsibility in virtual environments. As previously noted, one of the benefits of an open world is that the instructors can design and build learning experiences that are customized to the specific learning objectives of the class. These experiences can be created as part of a collaboration and can be shared among educators and students from anywhere in the world.

Weir also discusses the educational applications of the environment known as Whyville, which was designed for students to experience inquiry-based learning, such as the causes of epidemics, or the optimization of a manufacturing process. These experiences, while virtual, are authentic in that they allow the students to experience the effects of these situations on their avatars in the virtual space, and they can communicate with their peers and their instructors to share their thinking. Whyville as a somewhat simpler virtual world, is more accessible to students and instructors who may have less technical skills. On the other hand, Second Life, with its more complex graphics and design elements, has a higher barrier to entry for many educators.

Dickey (2005) conducted a qualitative case study of Active Worlds, a three-dimensional virtual environment like Second Life. This study explored two instances of virtual worlds being used for distance learning. In one case, the environment was used to facilitate asynchronous coursework, while the second case involved a class that met synchronously in the virtual world.

Active Worlds is a browser-based, three-dimensional virtual world that includes the key features described previously (communication tools, visually rich environments, and a virtual representation of the user as an avatar). One of the features highlighted in this study, which is also available in other virtual worlds, is the ability to shift perspective from first to third person. This option allows users to see the world from the perspective of their avatar, which places the user in the world, or to see their avatar as a character in the world, which is useful for certain tasks.

The first case study, a distance-learning business course at the University of Colorado, was examined using interviews and observational data. The introductory course is offered in traditional, online, and virtual world formats. The format of the virtual worlds course was a combination of asynchronous individual work and organized, synchronous group activities. The Active Worlds environment was utilized as a virtual learning space, in which students were able to navigate a variety of spaces, interact with multimedia objects, submit work, and receive feedback. For group assignments, the virtual environment provided a meeting place, where students could gather in virtual locations and use the communication tools to collaborate.

The findings of the case study suggest that setting the course in a virtual world offered distance learners with a sense of visual and physical "place," within which learning and collaboration could take place. Students who were physically separated in the real world could meet and work together in the virtual space, creating a sense of community. By creating familiar reference points, such as buildings and physical spaces, as well as visual representations of students and course content, Active Worlds provided “a *place* in which distributed learning is anchored in an environment that is both familiar and engaging” (p. 444).

In the second case, Active Worlds was used to create a meeting space for synchronous discussion and instruction for a 3D object-modeling university course. Students and the instructor logged into the Active Worlds University space at a predetermined time and “place” within the virtual world. The class was primarily managed by discussion using the text chat feature, as well as multimedia elements shared in the world or using the integrated web browser. One of the key features of virtual worlds that made this application interesting is that, in terms of 3D modeling, the students and instructors were meeting in a space constructed from 3D models. The instructor was able to refer to objects in the world, and to the underlying scripts, to demonstrate the principles of 3D modeling. This aspect of virtual worlds has a similar potential for the teaching of science concepts, in that 3D models of atoms or molecules can be created and observed from a variety of angles and perspectives. The results of the case study supported the concept that situated or authentic learning can be enhanced by the use of virtual worlds, when the students have a sense of presence, and the world creates a sense of space where students can engage with the content directly in ways that are not possible in the real world.

A second study by Dickey (2011) explores Murder on Grimm Isle, an immersive virtual game world designed as a tool for persuasive argumentation and presentation of evidence. The authors conducted a qualitative case study to gain insight into the motivating and engagement capacity of a narrative, game-based virtual environment for learning. The virtual world used in this study is an immersive environment that is navigated by the students. Unlike the open virtual worlds in the previous studies, Murder on Grimm Isle features a narrative structure that guides the learning. Students are presented with a murder mystery and asked to explore the environment in search of evidence. At the conclusion of the lesson, each student presents their conclusions, and must make supporting arguments based on the evidence.

The results of the study, based on direct observations, chat logs and interviews with students, found that students were curious about the design of the experience as a game, but that different game types might influence motivation differently for some students. In addition, students were able to construct reasonable arguments based on the evidence and information presented in the virtual world. However, the author notes that some students imposed their own meaning by making associations that were not explicit in the environment. This finding is particularly relevant when considering the use of virtual worlds for teaching science content. If students have different prior knowledge or existing misconceptions, then they could interpret or experience the environment in unintended ways that could hinder rather than enhance learning.

The final study in this section of the discussion is different from the previously discussed cases in two key characteristics. First, it followed a quasi-experimental design, combined with case study analysis. Second, it explored the possibilities for virtual worlds to create entirely new models for teaching and learning. The Schome Park Project (Twining, 2009) utilized the Teen Second Life open virtual world in a series of phases to test different experiences of students in different models of learning.

As demonstrated in several of the previously discussed studies, many examples of using virtual worlds for teaching and learning follow essentially the same model of education that exists in the real world. Students meet and interact in virtual spaces, receive instruction, and demonstrate their learning by completing and turning in assignments. Even within these virtual worlds, many institutions create replicas of actual buildings that exist on real campuses (Callaghan et al., 2009). Twining (2009) found similar tendencies when considering designs for their virtual world. Many of the visions included examples of the “like now only better” (p. 497) thinking that often surrounds the initial implementations of new technologies in education. Rather than recreating current educational models in virtual space, the study sought to consider entirely new possibilities that could be achieved using virtual reality.

The author notes two key elements of virtual worlds that made it ideal for this project. First, the flexibility of virtual worlds makes it possible to develop experiences that would be impractical or impossible in the real world. Second, when considering a variety of design elements, virtual worlds have the added benefit of being malleable in ways that the real world is not. It is possible in virtual reality to create, modify, discard, and reimagine spaces, objects, and entire worlds at little to no cost. This flexibility makes virtual reality an ideal environment for experimenting with new models for teaching and learning.

The project by Twining (2009) took place in three phases, with each phase transitioning from a more traditional to a more radical vision for what education might look like in a virtual world. In the first phase, the virtual space was modeled closely on a traditional educational environment, with buildings and classrooms devoted to specific tasks or content areas (i.e., science, sociology, etc.). In the second phase, students were given more authority to create the environment in ways that they felt best suited the needs of the learning objectives, while maintaining some basic elements that were designed by the instructors. In the third phase, there were two environments established. The first was conceived and created by the students, and the second was built by the instructors. The students were given total control over the new space and developed a game-themed environment for learning.

Also, each phase of the project involved more student control of activities and content to be studied. In the first phase, content adhered to traditional strands of learning (science, archeology, etc.) and was delivered in traditional instructor-led sessions in the virtual world. Phase two of the project featured more student led sessions and discussions, alongside the instructor-led sessions, following the same strands of learning. In the third phase, all traditional strands were led by students, and new strands were introduced either by student interest (time travel) or instructors (mathematics).

The Schome Park Project explored new and novel approaches to education, which were made possible by utilizing the virtual world of Second Life. Several key findings came out of the exploration, as supported by data collected via chat logs, interviews, and observations of a variety of media. First, it can be difficult for instructors and students to envision models of teaching and learning that break away from the traditional model. Second, once that traditional barrier is broken, virtual worlds can enable radically different educational models that are not possible in the real world. Since virtual worlds are not bound by the physical world and traditional models, entirely new paradigms for what is an educational environment can be explored. Third, and perhaps most importantly, virtual worlds offer the possibility to achieve deeper levels of experience. As Twining (2009) describes it, virtual worlds move from learning by doing to learning by becoming.

There are many examples of the use of commercial video games for learning. Key advantages of these applications are that they contain rich, often open virtual worlds, within which players have developed complex social and economic systems. Rather than create environments from scratch, teachers and students can use these existing structures to explore a wide range of concepts, from language and sociology, to epidemiology, to economics and statistics. The two cases discussed here involve the study of various aspects of World of Warcraft (Steinkuehler & Duncan, 2008) and Minecraft (Smaldone, Thomspon, Evans, & Voit, 2017).

In the first study, Smaldone et al. (2017) used the framework of the commercial game Minecraft for learning concepts of chemistry and chemical engineering in a higher education setting. Minecraft provides an open world, offering users the ability to harvest resources, build complex items, and create in a three-dimensional environment with relatively simple graphical interface and tools. The authors created a modified version of the game (a “mod”) that supported the mining, manufacture, and distillation of organic molecules and polymers. The study followed a qualitative case study approach, with performance data gathered based on the ability of participants to reach certain levels of achievement within the game.

Key elements and findings in the project included the concepts of differentiation, self-paced learning, gamification, and assessment. The items that could be crafted in the mod required a wide spectrum of knowledge of chemistry. For example, basic understanding of physical properties would lead users to harvesting rubber to create running shoes. More complex items could be crafted but required a more in-depth understanding of chemistry. It is important to note that all these items could be created by any user, as all the crafting recipes were available to all students. However, a student with more prior knowledge would be able to access the more advanced items more quickly. This levelling of tasks is one of the concepts that the authors mention in terms of assessment: the time and number of attempts required to achieve success.

The results of the study showed that participants gained an interest in and at least some enhanced understanding of organic chemistry concepts and processes. The study did not measure the effect of Minecraft as an intervention, but rather to gather feedback on the engagement and motivation of students to learn difficult concepts in a gamified virtual world.

Steinkuehler and Duncan (2008) explored the scientific nature of interactions in the massively multiplayer online roleplaying game World of Warcraft (WoW). The authors note that in virtual worlds such as the one in WoW, players interact, communicate, and collaborate to solve complex problems. This collaboration is a vital skill in the current and future educational model, and it takes place on a regular basis in virtual gaming environments. The authors used a qualitative analysis of discussions among players on online chat boards and reviewed the data in the context of "scientific habits of mind" (p. 530).

The analysis showed that there were several themes that would be considered forms of scientific discourse taking place on the discussion boards. For example, a considerable amount of discussion involved communicating for the purpose of mutual understanding or building upon prior knowledge to expand or refine ideas. Nearly a third of all discussions were coded as involving elements of scientific discourse, in the form of forming persuasive arguments, and supporting positions with evidence. There were examples of discussions involving concepts of reverse engineering, reflection, and scientific inquiry.

The study by Smaldone et al. (2017) was retrospective in nature and did not examine any performance or intervention data. Rather, the study demonstrated that, even without intervention from an educational perspective, virtual worlds are rich with opportunities to foster scientific thinking and informal inquiry-based practices. These worlds have the added benefit of having existing user-bases, which certainly include students in our current educational system. It is important to consider the opportunities to leverage these spaces for more explicit learning, as well as making connections to the practices that students are already exhibiting in them.

The studies in the previous two sections examined virtual worlds and environments that were not necessarily designed for the purpose of education, but which nevertheless offer frameworks that have been applied to create learning experiences across a wide range of content areas and educational settings. Shelton, Satwicz, and Caswell (2011) offer an analysis of educational games in the context of a variety of learning theories, and with a focus on the affordances available with established and emerging digital technologies. In this section, the focus will be on virtual environments and simulations that were purpose-built for teaching and learning of a variety of concepts. These tools offer opportunities that may not be possible or practical in the real world or provide options for control and data collection that are not readily available in traditional settings.

Two studies (Baser & Durmus, 2010; Finkelstein et al., 2005) explore the use of simulations for teaching and learning science concepts. The Physics Education Technology (PHeT) simulations created by the University of Colorado are interactive, web-based tools that allow students to experiment with a wide range of scientific concepts, from physics to chemistry and beyond. Users can design, build, and test different machines, chemical structures, or circuits. The advantages of the simulations are that students can interact with complex systems safely, and the digital nature of the simulations can make visible elements of the structures that would otherwise be abstract or unseen (i.e., flow of electrons, photons of light).

Baser and Durmus (2010) and Finkelstein et al. (2005) both studied the effect of using virtual laboratories for learning physics concepts related to the design and function of circuits. Baser and Durmus used a quasi-experimental design in which two groups of pre-service teachers learned about the concepts of circuits in one of two ways. The treatment group used the PHeT simulation, while the comparison group utilized real lab equipment to explore the same concepts. Finkelstein et al. (2005) used an experimental design to examine the same scenario, using PHeT.

The hypothesis in both studies was that students using the virtual laboratory equipment could develop an understanding of scientific concepts equally as well as students who had access to real laboratory equipment. The results supported the concept that learning with virtual laboratory equipment was equally as effective as learning with real laboratory equipment. Basur and Durmus (2010) found that, when controlling for prior knowledge, students in both groups performed equally well. The results in the study by Finkelstein et al. (2005) actually showed that students using the virtual laboratories outperformed the traditional laboratory students. This difference may be since the simulations allow for visualizations of concepts that are not available in real laboratory settings. This finding is particularly relevant to the current study, in which students will interact with an immersive virtual reality laboratory. While the level of immersion is different, the underlying concept remains the same.

Mei and Sheng (2011) designed, built, and tested a virtual reality simulation to assist medical students in their understanding of human organ systems. The authors developed a system that would allow medical students to examine and manipulate 3D representations of human organs. The environment also featured user avatars and interactions with instructors or doctors for access to course content. The authors surveyed medical students about their experience in the virtual learning environment and found that most students agreed that it was helpful to interact with the simulation of organs. Mei and Sheng also highlight that, by using the simulation for students to study organs, both safety and cost benefits were possible. This finding is relevant to the current study, which will save on the purchase and disposal of consumable chemicals and laboratory equipment, while giving students safe access to those chemicals and pieces of equipment in a virtual environment.

A study by Richards and Taylor (2015) used an experimental approach to evaluate whether there was any benefit to three-dimensional versus two-dimensional simulations in a biology course. The authors developed 2D and 3D models of the same concept, using NetLogo and Unity 3D, respectively. The Unity 3D engine is the application used in the current study, so its use here is of note. The authors used a pre- and post-test model of content knowledge on the topic of interest, to analyze whether there were any gains in 3D over 2D. Users were randomly assigned to the 2D or 3D environment. The results of the study generally showed that there was no significant benefit in using the 3D representation of animals in the model. However, it is important to note that this environment had only limited interactivity or immersion.

Padiotis and Mikropoulos (2010) used a virtual representation of an industrial milk production line to help university students improve their understanding of the operation and function of various components of the equipment involved in milk pasteurization. The authors used the structure of observed learning outcomes (SOLO) taxonomy to evaluate both the accuracy of students' responses as well as the level of learning demonstrated in a pre- and post-test model. Students had the ability to manipulate the equipment and adjust parameters of the system to observe the effect of their changes in the simulated factory environment. The authors note that most students in the study reported that they had a strong sense of being in the factory, and a perception that the equipment and processes they observed were very similar to real world equivalents. The results of the study showed that students exposed to the virtual environment were able to resolve misconceptions, describe the production process more effectively, explain the use and interaction of the various elements of the production line, and were able to gain knowledge of the internal working of the equipment, which might not have been possible in an actual factory setting.

One of the benefits of immersive virtual reality is that it allows for students to investigate and have authentic experiences in a safe environment that can be customized to the needs of the individual student (Kirkley & Kirkley, 2007). The level of immersion and the sense of presence can impact a learner's ability to engage with the learning experience (Scoresby & Shelton, 2011).

Ahn et al. (2015) tested the effects of immersive virtual reality versus a computer-based virtual world in the immediate and longer-term effects of a message on environmental consciousness. The study used a quasi-experimental design, with participants selected from a convenience sample at two different universities.

The virtual experience involved either the cutting down or growing of a tree in the forest. In the low-interactivity condition, participants watched the tree being cut down or grown over a 2-minute period, complete with visual and sound effects. In the interactive condition, participants used a haptic feedback rig to feel either the chainsaw cutting the tree or the pumping of nutrients to make the tree grow. Other than the haptic feedback, the sounds and visuals were the same as the low-interactivity condition. Results showed that, both immediately after delivery of the message and 1 week after the message delivery, the high interactive group experienced greater behavioral modification and environmental awareness (Ahn et al., 2015).

Passig (2015) used immersive virtual reality to examine the effect of VR on a series of different types of cognition. In the first series of studies, participants were faced with trying to understand difficult concepts. Examples included experiencing the world through the eyes and logic of a toddler and seeing the world through the eyes of a person with dyslexia. The studies used immersive experiences and compared their perceptions to them to control groups. In another series of experiments, the author tested a variety of learning needs, and showed that immersive virtual reality had the ability to produce improvements in inductive and deductive reasoning, and for aiding students with specific learning challenges. In each case, the immersive experience was superior to alternatives, including 2-dimensional virtual worlds. Scoresby and Shelton (2011) examined the relationship between visual perspective and the ability of users to achieve a state of presence or *flow*. They found that the level of engagement could be impacted by the perspective chosen for the environment. In the current study, the learning environment will be immersive, and take a first-person perspective.

Research Methods

As noted above, many of the studies conducted are qualitative in nature. In part, this tendency toward qualitative research may be due to the relatively new nature of immersive virtual reality technologies. There are simply not many examples or applications available for examination. The technical skill and knowledge required to develop these tools independently is still a barrier to entry, although more intuitive tools for content creation are becoming available, as VR technology begins to diffuse through different aspects of commercial, industrial, and educational arenas. Some studies, particularly those in the educational setting, followed quasi-experimental models, using samples of convenience, rather than strictly randomized assignments (Shadish, Cook, & Campbell, 2002; Steiner, Wroblewski, & Cook, 2009). However, some of the studies cited here followed true experimental designs, with randomized treatment and comparison groups. The method selected for the current study used utilize existing groups of students in different sections of a high school chemistry course. This method is in line with the previous research in the field.

Chapter 2 Summary

As noted, there have been several attempts to develop interventions to aid in the teaching and learning of chemical equilibrium (Banerjee, 1995). Many of the more recent studies have included application of technology and simulations to improve learning outcomes in this area. The challenges facing students and teachers alike range from issues of vocabulary to mathematical abilities, to the problematic nature of understanding interactions that cannot be observed directly (Annisa & Rohaeti, 2018; Cheung, 2009; Lucanus, 2011; and Maia & Justi, 2009). There is a need for additional opportunities for direct, hands-on interactions to develop a deeper understanding (Primastuti and Atun, 2018).

While physical limitations prevent students from working with and observing reactions on a molecular level, there is evidence that models and simulations can be effective when building understanding (Fogarty et al., 2012). Further evidence supports the use of scaffolding in activities to move students through a stepwise process of various representations of chemical equilibrium (Ghirardi et al., 2014). As noted previously, it is critical to bridge the gap between the abstract concepts and more concrete models or visualizations in chemistry (Maia & Justi, 2009). The use of analogies in teaching and learning of chemistry concepts has also been studied, and shown to be effective (Trey & Khan, 2007). Other studies demonstrate the effectiveness of allowing students to use inquiry and constructivist approaches to develop their understanding (Annisa & Rohaeti, 2018; Maia & Justi, 2009; and Ghirardi et al., 2014). All these factors point to the possibility that an interactive, virtual learning environment, such as the one developed for this study, may be well-suited to improving chemistry education as it relates to the concept of equilibrium.

The studies on virtual reality and simulations discussed here examined a wide range of virtual reality environments, from simple simulations to virtual worlds and immersive technologies. Rubio-Tamayo et al. (2017) describe a continuum of virtual reality, in which the distinction between the physical and the virtual is represented as a series of incremental steps from one end to the other. The current study would fall heavily toward the virtual end of the continuum. However, the purpose of these virtual environments is to enhance knowledge and skills that are applicable in the physical world.

While many of the studies focused on the experience of the user, rather than performance measures, those studies were illustrative of the engagement, motivation, and authenticity of learning that can be experienced using virtual reality. While some studies found that virtual reality did not show improvement in performance over real environments, this outcome is a useful result.

One possible impact of the current project was to enhance learning and provide deeper levels of understanding of science content, another key feature of the system was to allow students to have an authentic experience in a virtual environment that is completely safe and cost-effective. If virtual reality can provide the same degree of learning and transfer of knowledge, then these benefits of safety and cost would justify the adoption of immersive VR laboratories in science classrooms.

However, there is evidence in the literature to support the idea that immersive virtual reality can provide students with opportunities to learn that are superior to the experiences that they might have in a traditional educational setting (Dalgarno & Lee, 2010). Furthermore, as

Twining (2009) notes, the flexibility of virtual environments offers the ability to create learning experiences that are completely impossible to conceive in the real world. This combination of affordances could lead to entirely new paradigms for teaching and learning in virtual reality.

CHAPTER THREE: METHODOLOGY

Introduction

This study designed, built and tested an immersive virtual reality environment to support a student's understanding of the concepts of chemical equilibria and Le Chatelier's Principle. Each Student's knowledge was tested before and after their interaction with an immersive virtual reality environment. An analysis of student performance data assessed the effectiveness of the environment as a teaching and learning tool. Students also completed a post-intervention survey to assess the challenges that they encountered when using the intervention.

There is a large and growing body of research on the implementation of simulations, virtual reality, and immersive virtual worlds (Freina & Ott, 2015; Hew & Cheung, 2008; Passig, 2015). However, there is a gap in the literature on immersive virtual reality for teaching and learning of science concepts at the secondary level. Simulations exist for teaching and learning of chemistry, and there is evidence to support virtual environments as a learning tool (Dalgarno & Lee, 2010). Still, more research is needed to examine the implementation of immersive learning environment in specific content areas. This study contributes to the field by exploring the use of immersive virtual reality to support the study of chemistry at the secondary education level.

Many concepts in chemistry present a combination of challenges that include safety, cost, and phenomena that are impossible to observe directly (Bursztyn et al., 2017; Finkelstein et al., 2005; Mei & Sheng, 2011). Immersive virtual reality presents the opportunity to overcome all these obstacles, and this investigation will attempt to take advantage of those factors. This study brings the application of virtual reality to the secondary chemistry classroom for the teaching and learning of the concepts of chemical equilibrium.

Research Methodology

The proposed study followed an educational design research (EDR) model to explore the potential or immersive virtual reality for improving learning outcomes for high school chemistry students. The limitations of the classroom environment make a true experimental design impractical (Campbell & Stanley, 1966), therefore the study employed EDR, which has a foundation in design-based research and has been developed and structured to provide practical knowledge on real educational settings (McKenney & Reeves, 2012). The design is described in more detail in the Educational Research Design section. The EDR model was selected for the current study based on the constraints of the classroom, the complexity of the learning problem being studied, the research questions, and the nature of the proposed intervention. The study included participants from multiple sections of a college preparatory high school chemistry course.

Statement of the Problem

High school chemistry provides students with foundational knowledge of physical science and the curriculum includes many concepts that are required by both state and federal standards. This learning objective is complicated by the nature of atomic and molecular interactions, which have proven difficult to reify. Among the major concepts covered in chemistry, chemical equilibria and Le Chatelier's Principle are prominently described in the descriptions of the Next Generation Science Standards. These concepts present significant challenges for both students and teachers (Ghirardi et al., 2014).

These conflicting demands, combined with decades of research devoted to improving understanding of this topic establish a distinct need for new interventions. The current study designed, developed, and tested one possible intervention utilizing the emerging technologies of immersive virtual reality and intelligent tutoring systems. Educational research and practice have begun to shift toward experiential and inquiry-based learning in recent years (Freeman et al., 2017; Glowa & Goodell, 2016). However, in the study of chemistry, especially at the secondary level, there are limitations on experiential learning, in terms of safety, cost, and practicality, that all lend themselves to working in a virtual environment (Bursztyn et al., 2017; Finkelstein et al., 2005; Mei & Sheng, 2011). Students in a virtual world can experiment freely with a wide range of substances that would be hazardous or in a real laboratory setting, and can repeat those experiments as often as necessary without consideration of cost of materials or equipment.

Prototype Development and Research Questions

Educators employ multiple approaches to teaching abstract chemistry concepts, such as equilibrium. Some of these approaches involve conducting hands-on activities involving reactions in equilibrium, the use of analogies and models, and the use of technology and simulations. Each of these approaches has advantages and drawbacks, as noted previously.

The current study combined elements of all three approaches (hands-on, analogies, and technology), taking advantage of the accommodations afforded by virtual environments. The study had two main objectives. First, to design, develop, and test a prototype of an immersive virtual reality learning environment for enhancing understanding of chemical equilibrium. This prototype was created using commercially available hardware and software, but the environment itself, the design of the learning activities, and the interactions within the environment were developed and built as part of this project. The prototype was then pilot tested by a group of high school chemistry students. The second objective was to collect data during the pilot testing to assess student performance on a content-based pre-test and post-test, as well as data on user experience. The analysis of this data provides guidance for future iterations in the development of the learning environment.

The data collection and analysis were guided by three questions, as previously stated, and summarized here:

Research Question 1: To what extent do high school chemistry students learn concepts of chemical equilibrium via the implementation of an immersive virtual reality environment?

Research Question 2: What technical and conceptual challenges do students encounter while using a virtual environment for learning chemical equilibrium concepts?

Research Question 3: What aspects of usability and user experience correlate to student performance?

Given the research environment, the research methodology selected was educational design research (EDR), which is described in detail Educational Design Research section. The current study consisted of a single meso-cycle of EDR, including analysis, design, implementation, and evaluation of a prototype learning environment. The primary data collected to address Research Question 1 was in the form of pre- and post-test scores among a group of high school chemistry students with similar academic backgrounds. Data related to the second research question was collected via a post-intervention survey. Both the pre- post-test data and the survey data was used to evaluate the prototype and guide future iterations of the learning environment.

Educational Design Research

Educational Research Design (EDR) is a model that is cited as offering practical knowledge in real-world settings for education (Van den Akker et al., 2006). The EDR model is a type of design-based research that is iterative, comprising multiple stages in which an intervention, policy, or teaching model is developed. The EDR model includes several phases, or micro cycles, each of which represents a major component of the research project (Reeves & Amiel, 2008). There are a total of four phases in the EDR model. Three phases, called micro-cycles, comprise the fourth phase, which is a single meso-cycle. These phases can then be repeated as necessary for subsequent iterations of the project. The first phase is the analysis and exploration phase, during which the research problem is identified, and prior knowledge is assessed. The first two chapters of this dissertation served as the analysis and exploration micro cycle. The next two micro cycles are design and construction and evaluation and reflection. These microcycles were be carried out during the current study. These three micro cycles, taken together, represent a single mesocycle of the EDR model (McKenney & Reeves, 2012). Completion of the project represents the completion of one meso-cycle (see Figure 1). The data gathered during the study will be used to guide future development of the learning environment, which will then be tested through further meso-cycles in future years.

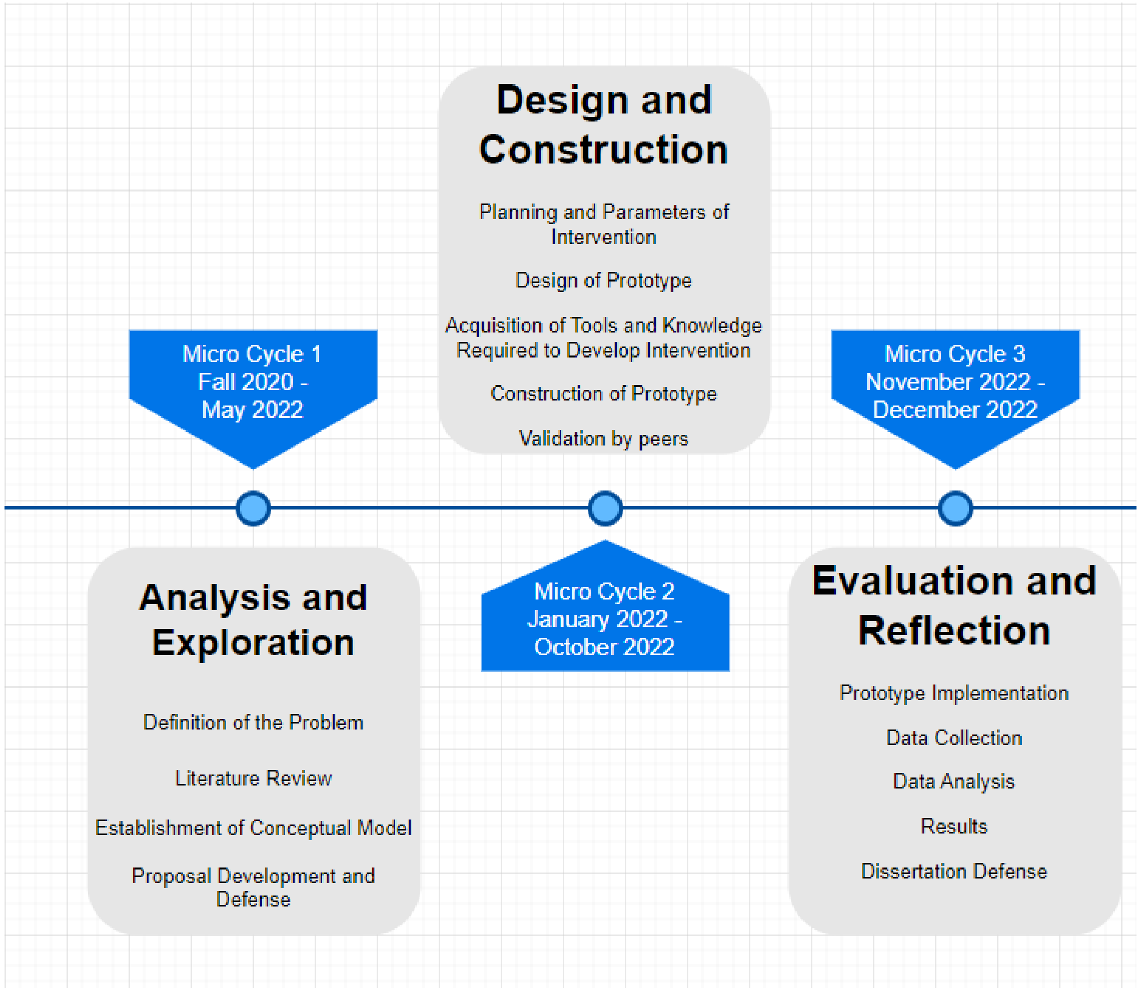

Figure 1 Overview of Education Design Research Cycle

Micro Cycle 1: Analysis and Exploration

The results of the first analysis and exploration micro cycles are included in the Introduction and Literature Review Chapters of this dissertation. In those chapters, the research problem was identified and the challenges related to the problem were outlined. The teaching and learning of equilibrium and Le Chatelier's Principle are long-standing challenges for educators and students at all levels. While many methods and tools have been developed, the issue remains unresolved. The prototype intervention developed for this study targets the problem by applying an immersive virtual reality learning environment to the teaching of chemical equilibrium in a high school chemistry class. As described in the previous chapters, there are many studies on the effective teaching and learning of equilibrium, and a growing body of research on the use of virtual reality in education. There has not yet, however, been a study that applies virtual reality to chemical equilibrium in a high school education setting.

Micro Cycle 2: Design and Construction.

The second micro cycle for the current study included the design and construction of a virtual reality learning environment for the teaching and learning of equilibrium concepts. The outline of this design and details of its construction are described in greater detail later in this chapter. The prototype for the intervention will use currently available virtual reality hardware and off-the-shelf design software to build an immersive, interactive environment for students to explore. Within the environment, students are guided through a series of challenges related to chemical equilibrium. In order to successfully complete each challenge, students need to apply the knowledge gained in previous challenges. The learning environment includes multiple pathways, allowing students to repeat levels, if necessary, to encourage mastery of each concept.

The prototype consists of a series of locked rooms that students must unlock to advance in the environment. Each room was designed and mapped to target a specific question within the assessment. Taking a constructivist approach, rooms present the concepts of chemical equilibrium first as a series of analogues, and later applying the same concepts to actual chemical reactions. The environment also features a branched structure, so that students who fail to successfully solve the puzzle after a prescribed number of attempts will be directed to another room with a similar challenge. A crosswalk chart was developed, linking each room to the assessment question for which it is constructed. This information may be utilized in the design of future iterations for certain assessment questions that are shown to me more or less successfully answered following the intervention. Prior to the pilot implementation in the classroom, the prototype was evaluated by a high school science teacher to ensure that the design of the puzzles matched the intended concepts.

In addition to the design and construction of the learning environment itself, this micro cycle included the design of a pre- and post-test assessment to evaluate students' understanding of equilibrium concepts. The pre-test was administered prior to the intervention, at which time students will have had prior teaching on basic concepts of chemical reactions. They will be familiar with the concepts of chemical changes, reactant and products, as well as endothermic and exothermic reactions. At the time of the pre-test, the students have had no direct instruction on chemical equilibrium. The virtual reality learning environment represented the only instruction that the students receive. Following the intervention, students took the post-test. Any change in performance from the pre-test to the post-test should be attributable to the intervention.

Micro Cycle 3: Evaluation and Reflection

The final phase of the meso-cycle included evaluation and reflection. In the case of the current project, this evaluation and reflection included analysis of the results of the pre- and post-test, along with analysis of the Virtual Reality Neurological Questionnaire, System Usability Survey, and the User Reaction Survey. The pre/post-test analysis was used to evaluate whether students improved significantly in their performance, which would support the effectiveness of the intervention as a teaching and learning tool. If students fail to perform well on any specific elements of the content, that may also be instructive as to subtopics that were not sufficiently represented in the learning environment. The results of the user experience surveys were evaluated to determine which aspects of the learning environment were effective from a user experience perspective. All this information may be carried forward into the analysis and exploration micro cycle of the next meso-cycle if future iterations of the learning environment are developed.

Justification of Research Model

Educational design research is best suited to problems that do not have a clearly defined, pre-existing solution (Plomp, 2009). As described in the prior two chapters and summarized above, there is a large body of research on the teaching and learning of chemical equilibrium and Le Chatelier’s Principle. These concepts are included in many state and federal curriculum standards (Massachusetts Department of Elementary and Secondary Education, 2016; NGSS Lead States, 2013). Smith and Gilbert (2019) followed an EDR model to develop and evaluate a gesture-based technology intervention for teaching and learning mathematics. Data gathered from student performance and usability feedback were discussed as guidance for future development of the prototype.

A study by Tare, Shell, and Jackson (2020) tracked student use of certain features in an online literacy platform, and matched performance to feature usage to gather evidence on the aspects of the platform that provided the most engagement and support. A recent study by Morch (2020) followed a design research model to test formal and practical learning opportunities in a virtual world environment. The author found that a combination of domain specific (content) knowledge and existing skills (games) could be used to help bridge gaps in learning. These studies did not involve immersive virtual environments but demonstrate the use of EDR to test the effectiveness of specific educational technology interventions in educational settings.

Research Design

The current study was carried out in five sections of a high school chemistry course. All five sections were taught by the same instructor, and the students participating in the study represented a heterogeneous mix of prior knowledge and abilities. The sections were all of the same, college preparatory academic level, which includes students from a broad range of achievement levels in mathematics and science. This group of students, compared to honors or advanced placement level students, typically struggle to grasp the concepts of chemical equilibrium.

The data collection followed pre- and post-test model, with assessments compiled based on state and federal standards for high school chemistry. The exact questions that comprise the assessments are provided in Appendix B. Students received the intervention utilizing the immersive virtual learning environment described below. Following the delivery of the intervention to all participants, a post-test was administered. The result from the pre- and post-test performance were analyzed to evaluate the effectiveness of the virtual reality intervention as a learning tool. In addition, students completed a user experience survey, including an abbreviated version of the Virtual Reality Neuroscience Questionnaire (VRNQ), the System Usability Survey, and items from the User Reaction Survey. Questions from these survey instruments gather information concerning user experience and usability of virtual reality environments. Respondents provide feedback on a Likert-type scale. Results of the survey were analyzed in relation to the second research question and will be used to plan future iterations of the learning environment.

The educational design research model has been described previously in this chapter. The key findings include the performance of the students on the pre- and post-tests. Additional data on the usability of the prototype was be collected via a post-intervention survey. This data was analyzed to evaluate the success of the prototype as a teaching and learning tool, and to provide guidance for improvements to future iterations of the intervention.

Learning System

The prototype learning system incorporates immersive virtual reality and a rudimentary intelligent tutoring system in an environment that adapts to the skill level of the learner as they progress through a series of tasks. The learning objective of the project is related to the Next Generation Science Standards (NGSS Lead States, 2013) for chemical reactions, specifically, those concerning the concepts of chemical equilibrium and Le Chatelier's Principle. In the environment, students complete a series of tasks, with increasing complexity, that require a mastery of the fundamental concepts involved. There are several underlying concepts that need to be addressed, including the ability to interpret a chemical equation, exo- and endothermic reactions, the relationships between reactants and products, equilibrium, and factors that can upset equilibrium.

Southgate and Smith (2017) highlight several key considerations that served as important guidelines in the planning and development of this project. First, the empowerment of the teacher was ensured by the fact that the lead investigator will be the instructor of the course under study. The time and resources required to build and pilot this system will be significant and estimating the need for resources was vital in the success of the project. Duty of care is a legal concept that is familiar to the instructor, who teaches laboratory science courses, which follow strict safety protocols to protect both the student and the instructors. It was also important to be aware of and stay current with research into the effects of immersive VR on the development of young people (Southgate, Smith, & Scevak, 2017).

As with any new technology, especially those that allow communication in computer-mediated environments, it is important to ensure that children are protected from potential harmful material or interactions. The current project did not have connections to the Internet, and care was taken to establish norms of safety and respect among individuals who use the system in a collaborative manner. The project also utilized minimal emotional stimuli, beyond the potential stress of solving a problem at hand.

The format of the project followed a gamified model, in which the students first complete a series of training tasks. Games with simple constraints balanced with user autonomy can be effective learning tools, as demonstrated by both educational and commercial game design (Shelton, Satwicz, & Caswell, 2011). Each training task was designed to measure the user's proficiency in the various concepts outlined above. If the user completes a predetermined number of challenges correctly, then they progress to the next task. If the user fails to complete the tasks, then they are guided through remedial activities to assist them with each task, as required. In future iterations of the game, the final task, and the upper limit of difficulty, can be modified for different levels of the course, with extension activities designed as bonus levels, unlocked by successfully completing certain achievements. A basic storyboard for the progression of tasks is shown in Figure 2, while a more detailed description of the branched structure will be provided in the following sections. The gamification elements of the project are primarily structural (i.e., tasks as levels and achievement unlocks of special activities), and a scoring system. However, future versions of the project could incorporate additional aspects of gamification, including a narrative structure, collaboration among players, and experience levels, as well as leaderboards. A more detailed description of the prototype is provided in the next sections.

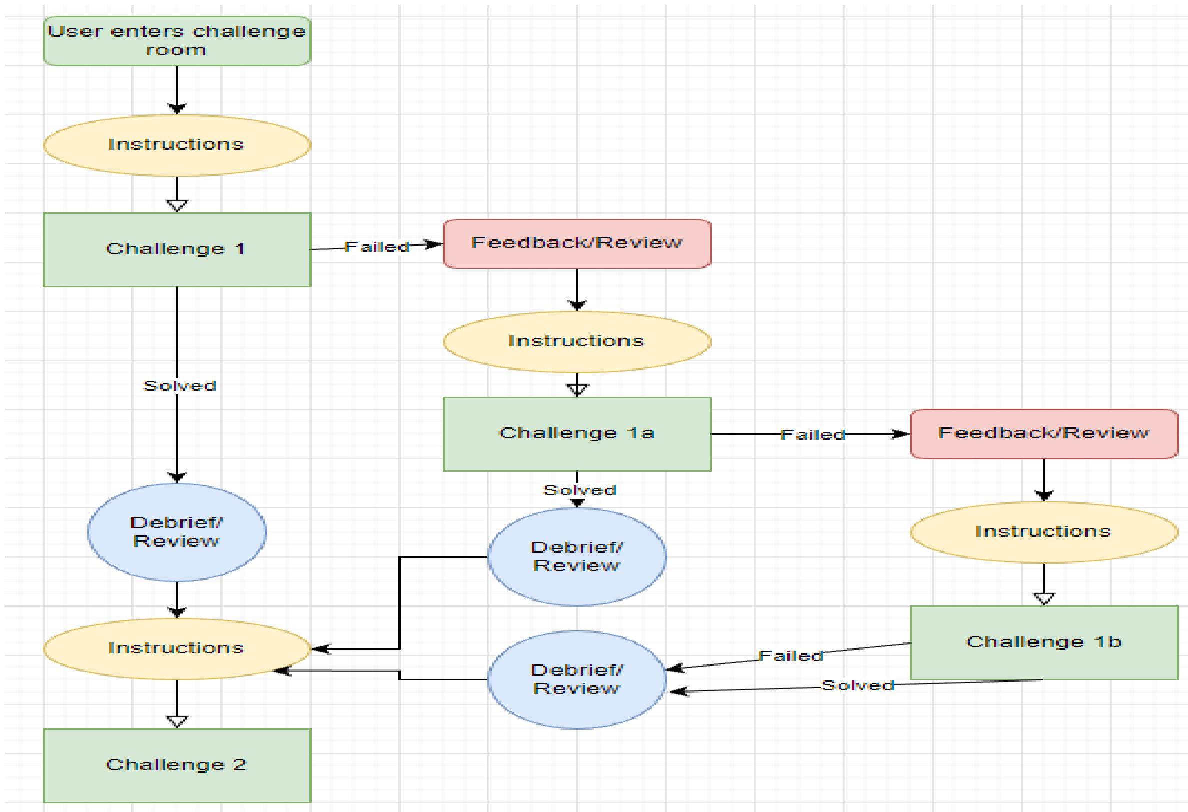

Figure 2 Sample task progression for the prototype

Intervention Timeline

The intervention will take place during a combination of 60-minute class periods and scheduled appointments during flexible study periods for each group. In the first session, participants were given a brief overview of the VR hardware, safety briefing, and explanation of the key types of actions that would be required in the prototype environment. They also completed the pre-test assessment. Students were not provided with their scores or the correct answers following the pre-test to avoid issues with memorizing the correct answers for the post-test.

During the second session, participants used the VR systems to engage in the learning activities outlined in the following sections. While some students may progress through the activities more quickly than others, a full class period or study period will be devoted to using the learning environment.

In the third session, participants will complete the post-intervention assessment and the usability survey. These assessments did not require the full class period, and any remaining time was utilized for regular daily classroom activities. The outline of the intervention timeline is presented in Figure 3 below.

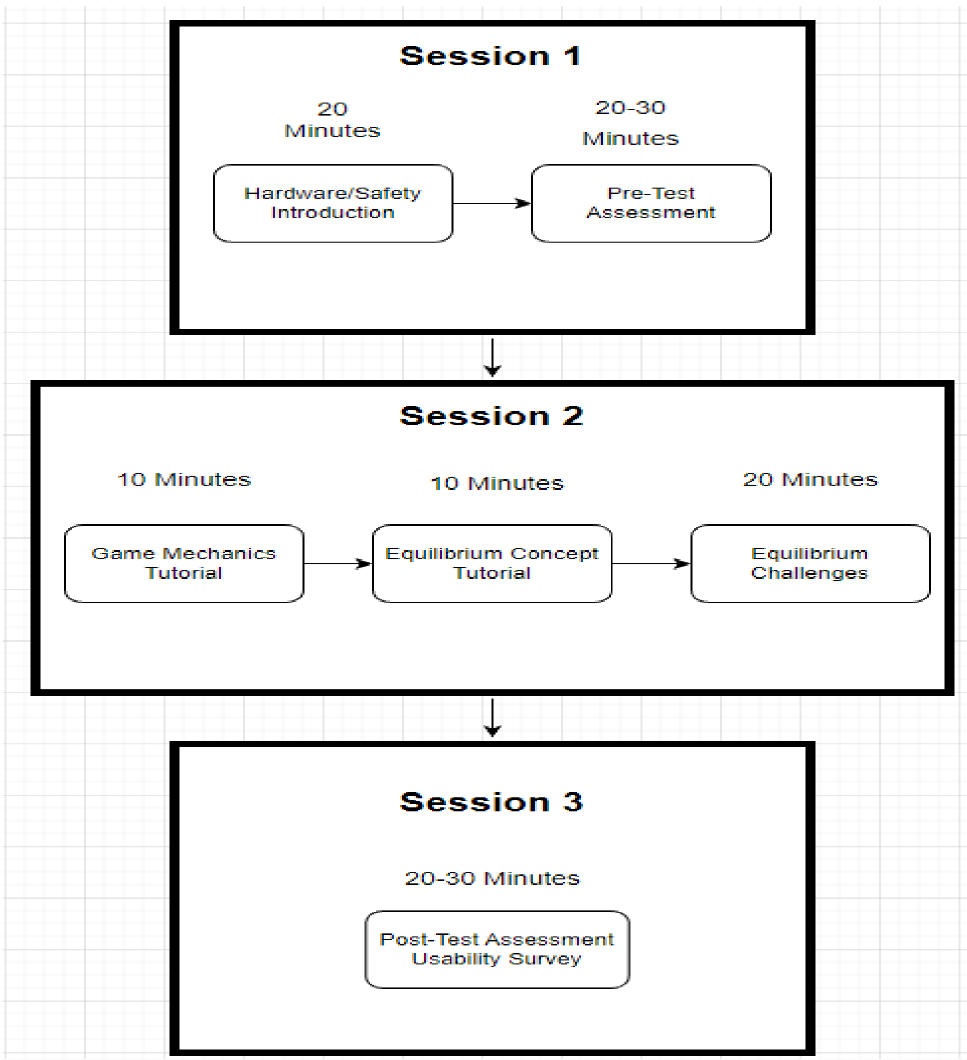

Figure 3 Intervention Timeline

The daily schedule of classes at the research site follows a "waterfall" rotating schedule of 7 blocks, in which 5 classes meet each day and two blocks are dropped. Based on this schedule, each of the 5 class sections participating in the study will have at least one day in which their block is dropped. Therefore, the time required to ensure that each of the 5 sections meets 3 times was approximately 5 school days, with at least two of the sessions being on non-consecutive days.

Prototype Design

As noted above, the planned prototype utilized a gamified structure to encourage students to explore scenarios designed to build knowledge of equilibrium and Le Chatelier's Principle. In this section, a more detailed description of the learning environment is provided. Images from the completed prototype and a detailed description of specific elements of the learning environment used in the classroom are provided.

Design Considerations

There are several affordances of virtual reality that make it a potentially powerful tool for constructionist learning models (Girvan & Savage, 2019). The proposed prototype will be designed to take advantage of these affordances as they relate to teaching and learning of chemical equilibrium concepts. As noted by Shelton and Scoresby (2011), aligning game design with learning goals is a crucial aspect in developing game-based learning environments. The proposed project will maintain this alignment by directly modeling specific challenges within the prototype to assessment questions based on curriculum standards.

First, the design will take advantage of the immersive nature of virtual reality to create an environment in which the learner is able to simulate a hands-on experience (Hanson & Shelton, 2008) The immersive nature of VR provides a sense of presence for the learner in an environment that can be designed and altered by the educator in ways that a traditional classroom environment cannot (Dass, Dabbagh, & Clark, 2011). This experiential learning is a key component of constructionism, and the proposed prototype will be designed to promote the learner's perception of actively experiencing the learning activities. The combination of active participation and autonomy with guidance and scaffolding are important components of a successful immersive learning environment (Shelton & Hedley, 2003). The structure of the proposed environment will be designed with the expected outcome of appropriately balancing these factors.

A second affordance of VR that will guide the design of the prototype related to constructionism is flexibility (Girvan & Savage, 2019). This flexibility is one of the key aspects of the intelligent tutoring system in the planned prototype, and it is one of the strategies that is the most difficult to replicate in a traditional classroom. While it is difficult and often impractical to generate multiple activities for students, especially considering the safety and supply issues involved in conducting multiple reactions simultaneously, the opportunity to allow students to explore different paths will be a central component of the prototype. This branching is also a consideration in designing the prototype within a constructionist framework. As noted by Kynigos and Futschek (2015), an environment that presents learners with opportunities to explore and test ill-formed problems, called a microworld, is a fundamental component of constructionism. The proposed intervention will represent a microworld in this sense.

Hardware

The hardware used by the learners in this project were Meta Quest 2 standalone virtual reality systems. This hardware system was selected based on several key considerations. The platform includes both head-mounted displays (HMDs) and handheld controllers that allow users to manipulate objects in the learning environment. It is critical for the proposed project that the learner can manipulate and interact with items within the virtual space in a way that is analogous to the use of real-world equipment wherever possible. As noted previously, the ability for the user to act upon and have a perceived influence on the environment is a key factor in the level of immersion experienced. The Meta Quest 2 systems also offer motion tracking at both a room scale and a stationary scale, creating opportunities for players to move in a variety of ways. wireless mouse with tracking capabilities. Using a standalone system reduces the overall cost and logistical challenge of being tethered to an expensive computer system. Finally, the Meta Quest 2 hardware is compatible with the development software described below.

<u>Software</u>

For the immersive virtual world, the project will be programmed using the Unity3D software suite (Figure 4). Unity is an industry-leading set of tools that allow for the creation and execution of applications across a variety of platforms (Goldstone, 2011). Unity is compatible with all major virtual reality platforms, and offers the ability to tailor features of the application to the specifications of the Meta Quest 2 system. Recent updates to the Unity platform include built-in support for virtual reality applications that made the development of the prototype a more streamlined process. The environment and interactive elements of the project will be created in Unity. The majority of the structural components of the environment (e.g., walls, floors, doors) were obtained from a freely available set of assets on the Unity Store. These components were combined and modified to create a customized environment for the prototype. Other assets within the environment (e.g., audio clips, particles representing molecules, signage, teleportation landmarks), were created by the researcher by modifying primitive objects available within the Unity platform. All of the coding for the operation of the environment was developed by the researcher using C# programming language in Microsoft Visual Studio, which is native to the Unity development environment. The cross-platform nature of Unity3D will allow for easy porting to other VR systems if necessary for future iterations of the intervention.

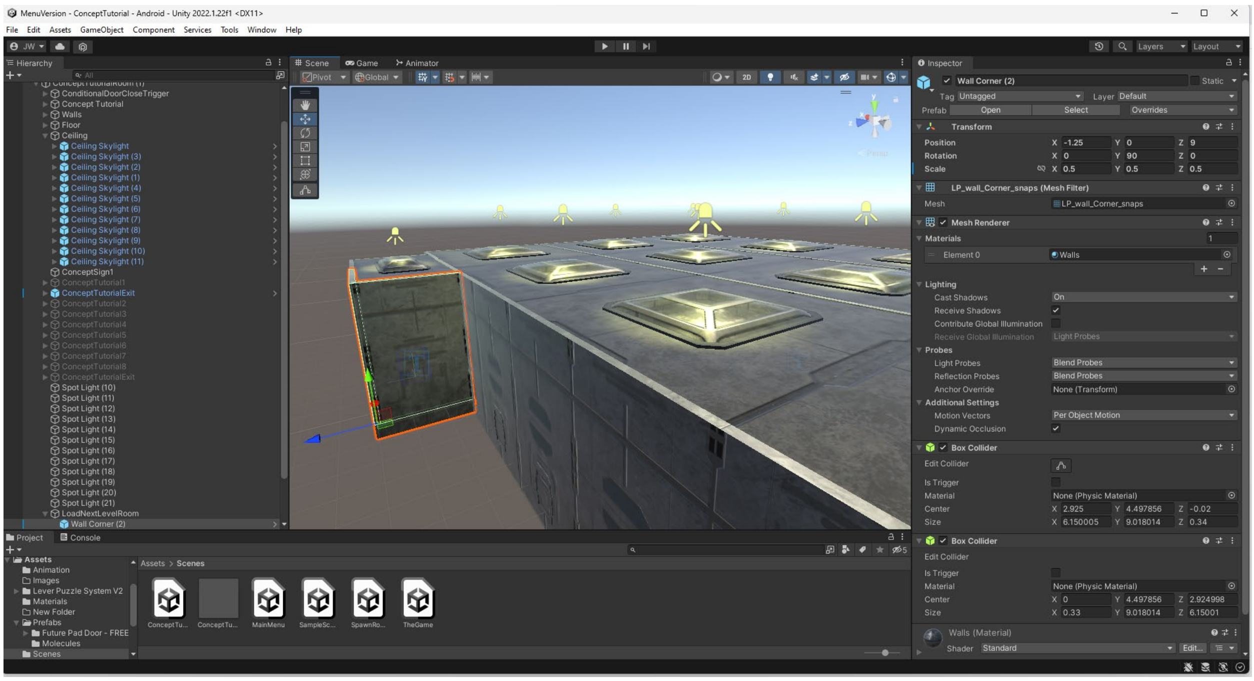

Figure 4 Unity 3D Integrated Development Environment

The ITS system consisted of a combination of structural features, such as physical pathways through the environment (Figure 5), remedial feedback following unsuccessful completion of a challenge, and multiple versions of challenges with similar concepts. All of these task management and task progression elements were developed by the researcher and integrated into the learning environment through the design of the environment itself and programming developed directly in Unity3D.

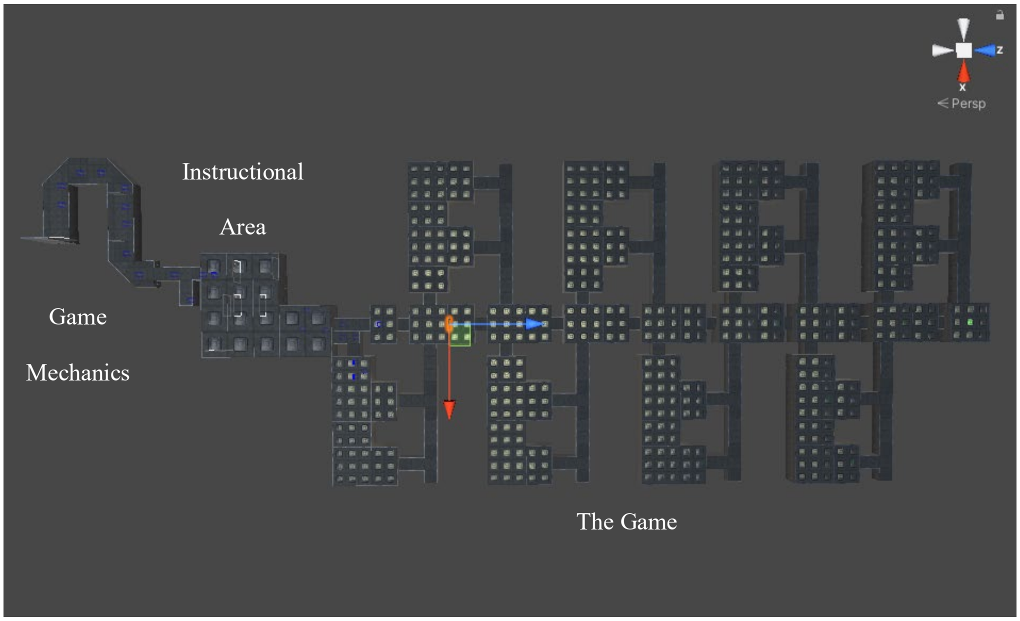

Figure 5 Learning Environment Structure

The structure of the environment shown above in Figure 5 shows the three main sections of the environment described in more detail below. The curved section to the left contains the tutorial on game mechanics. The rectangular section to the right of the game mechanics tutorial area includes the instructional area where players learn about concepts of chemical equilibrium. The largest section of the structure represents the learning game itself. The straight center section leading left to right in the image represents the main pathway that players follow to access challenges. The sections that are staggered toward the top and bottom of the image represent the secondary pathways that players take if they incorrectly answer one of the main challenges. Each of these secondary pathways includes two different exit points, depending on whether the player solves the first or second alternate challenge correctly. The final room to the right of the image displays each player's final score and the pathway to exit from the game.

Learning Environment Design

The following sections describe the process utilized to develop the prototype. Details of the development methodology used to manage the process are included. A description of the hardware and software used to develop and test the prototype is provided. Finally, a detailed description of the prototype learning environment is included to demonstrate the key features of the environment itself.

Agile Methodology

The process for developing the prototype followed the Agile Scrum software development model. Agile is a term used to describe a range of project management methodologies follow a set of principles that was developed in the software industry in 2001 (Waqar, 2020). Specifically, this project was developed using the Scrum version of Agile. While there are many development models used in software development, Agile Scrum was selected for this project. Scrum was selected for its flexibility and because it follows a three-phase model similar to the overall design of this study using the EDR Model described in the Educational Design Research Section (Mercan & Becerikli, 2020). The first phase of Scrum is planning, where the key requirements of the project are established and defined. The planning phase may be revisited if features need to be revised following the development process. The main focus of the planning phase is to establish the user requirements and the specific tasks that will be required to meet those needs. These tasks are organized using a Scrum board, which is a project management tool that displays tasks as individual items that can be in the Backlog (not yet started), In Progress, of Completed. In the current project, the Scrum board consisted of individual sticky notes placed on a wall. Figure 6 shows the Scrum board at several points during the development process, as tasks were added, revised, initiated, and completed.

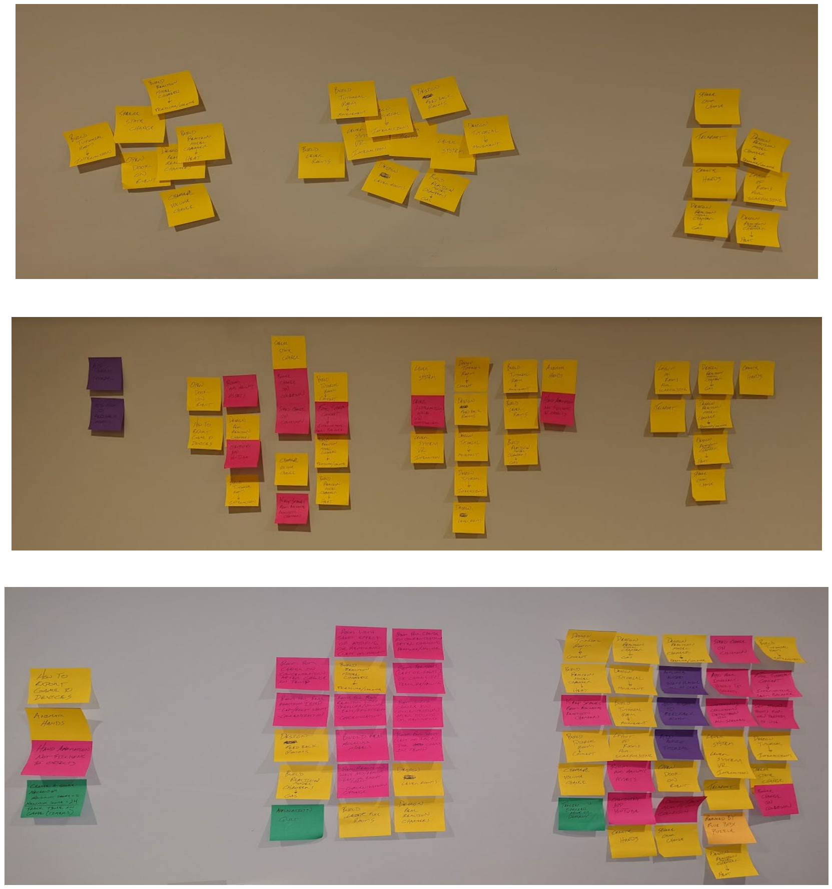

Figure 6 Scrum Board

The second phase is the development phase, which includes an iterative process of developing features, testing, evaluating, and revising. This process is made up of a series of Sprints, which are short timeframes during which specific tasks must be completed. Each task is evaluated regularly and sprint timeframes are modified as needed. At any point in the Development phase, it may be necessary to return to the Preparation phase when new requirements are added or new tasks become necessary. In the case of the current study, there were several iterations of the development phase that necessitated additional time spent in the planning phase. This iterative process occurred as features of the prototype were being developed that either required additional game elements be developed first, when specific features were found to be impractical, or when more efficient methods of achieving specified objectives were discovered. This phase also included the peer testing and updates based on peer feedback. The new or revised tasks were added to the Scrum board using different colored sticky notes to represent different phases of the development process (Figure 6).

The final phase of the development process is the Close phase, which includes the final distribution of the software. In the case of the current study, this phase was represented by the implementation of the prototype in the classroom with actual students.

Learning Environment: Main Menu

One of the issues discovered during initial testing and peer testing was that the three sections of the environment described below were integrated in such a way that a player who had to stop for any reason during the game had to complete all three sections in order when they resumed using the prototype. A modification was made to the application to introduce a Start Menu (Figure 7), which allowed players to select from one of three entry points: a) the beginning of the game (Game Mechanics Tutorial), b) the instructional area (equilibrium concept tutorial), or the game (a series of challenges related to equilibrium).

Figure 7 Start Menu Options

Learning Environment Part 1: Game Mechanics Tutorial

Students access the learning environment using Meta Quest 2 virtual reality system. The system consists of a standalone head mounted display (HMD), and a pair of handheld controllers. The HMD and the controllers are tracked in real time allowing the students to move in and interact with the environment directly. Upon entering the environment, players are placed in a corridor with instructions on how to use the teleport system. White markers on the floor indicate locations that players may select to move through the game. Both text and audio guidance are provided to direct players through this section of the environment. After moving through a corridor, players are introduced to other key aspects of the game, such as various types of doors and manipulating objects within the game to trigger actions. As each new game feature is introduced, players receive both audio and visual instructions (Figure 8)

In the example below, players are introduced to the game mechanic of manipulating objects to activate doors. The cylinder on the left side -labeled “Grab This” must be moved onto the right-hand box – labeled "Place Here" in order to open the Action Based Door. In addition to the visual cues, an audio explanation plays when the player moves to the teleportation point labeled "Stop Here."

Figure 8 Game Mechanics Tutorial Example

As students advance through the Games Mechanics Tutorial, they learn how to navigate in the environment and interact with objects. These skills will be used throughout the rest of the environment, in combination with the information provided in the Equilibrium Concepts Tutorial described in the Learning Environment Part 2 section. When players complete this portion of the learning environment, they will advance to the Equilibrium Concepts Tutorial. Once the Game Mechanics Tutorial is completed, students can skip this section via the Start Menu in future sessions as needed.

Learning Environment Part 2: Equilibrium Concepts Tutorial

The second part of the learning environment provides audio and visual content instruction on the concepts of chemical equilibrium. The audio instruction and interactive models in this section comprise a self-contained lesson on chemical equilibrium. The only prior knowledge assumed is a basic definition of a chemical reaction. Players move through a series of stations at which they learn about reversible reactions, equilibrium, and changes to systems at equilibrium (Figure 9). Each type of reaction change is correlated to one of the change types that are demonstrated in the assessment questions. These changes are then correlated to the different challenges that will be presented in the third section of the learning environment.

This section also allows players to further practice the movement and manipulation game mechanics that were presented in the game mechanics tutorial and that will be critical to complete the game in the third section. No additional external teaching is provided to students outside of this learning environment. However, the Game described in the Learning Environment Part 3 section includes additional concept reinforcement in the form of audio explanations, text descriptions, and visual representation of reactions.

Figure 9 Concept Tutorial Example

Learning Environment Part 3: The Game

The overall template for the learning environment was a breakout room. Breakout rooms consist of a series of tasks and activities that participants must complete to "escape" from the room. In the case of this learning environment, the challenges require participants to manipulate a series of puzzles to unlock each door in a series of rooms. In order to advance through the environment, students must solve the challenge correctly. If they make the wrong choice, they must take a different path through additional rooms, where they receive further instruction. The puzzles in each room present different components that were designed as analogues to chemical systems at equilibrium. The analogues become more complex and incorporate the key factors that influence systems at equilibrium. Players are given instructions on how to open these doors through the Game Mechanics Tutorial, and each room in The Game section of the learning environment includes an audio prompt to move in a given direction when a new door is opened.

Prior to entering the learning environment, students are told that they will receive a score at the end of the game. This score is recorded as part of the post-test assessment. Each player begins the game with a score of 25. Each time a student fails to solve a puzzle, a point is deducted from their score. The maximum number of challenges that a student could attempt is 24. Students who solve each challenge on the first attempt receive a perfect score of 25. Students who answer every challenge incorrectly would receive a score of 1. The final score is displayed to the student when they reach the final room in the learning environment.

Each analogue puzzle consists of a reaction chamber that contains spheres of various colors that move randomly within the chamber. When two spheres collide, code attached to each sphere is triggered to determine whether a "reaction" should take place. If the colliding spheres are of the correct color combination, those spheres are destroyed and two new spheres of different colors are created. This behavior by the spheres represents a reaction taking place and models the interaction of real atoms and molecules in a chemical reaction. Players are challenged to identify specific changes that could be made to conditions in the chamber in order to produce a prescribed result. Colored cylinders are selected by the player, using their virtual hands, and placed into a corresponding slot to identify their choice (Figure 10). Each challenge requires the player to select two correct answers. This choice was made to minimize the likelihood of a player correctly solving a challenge with a single random guess. In the event that a player makes one correct and one incorrect choice, an audio prompt advises them that they should change one of their choices. The additional audio prompt was added as the result of feedback from the peer testing process. Players have the option to add and remove choices, as long as they do not have two incorrect choices at any given time. Making two correct choices prompts a door to open, leading to a debriefing room where players see the results of the changes that they made and hear an audio explanation of those changes (Figure 11)

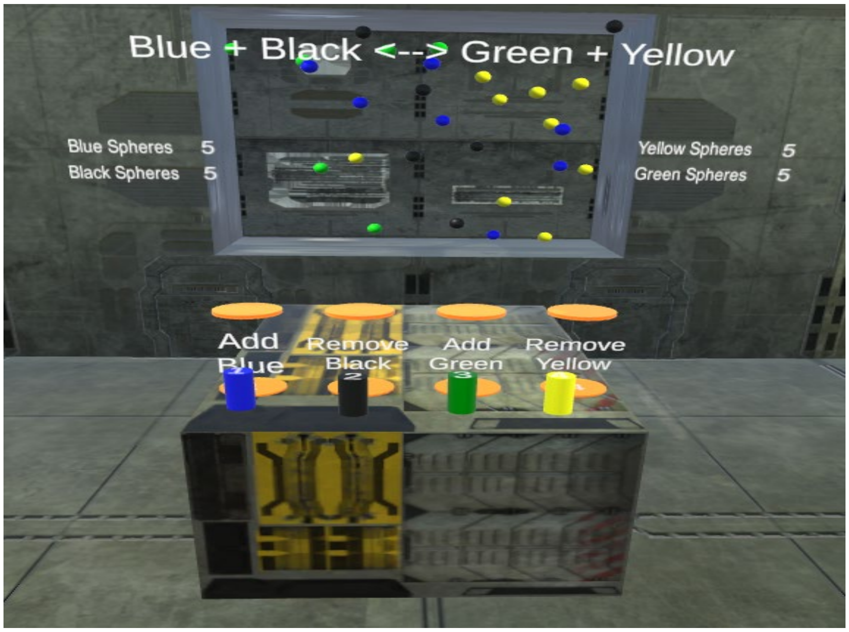

Figure 10 Example of Analogue Reaction Puzzle

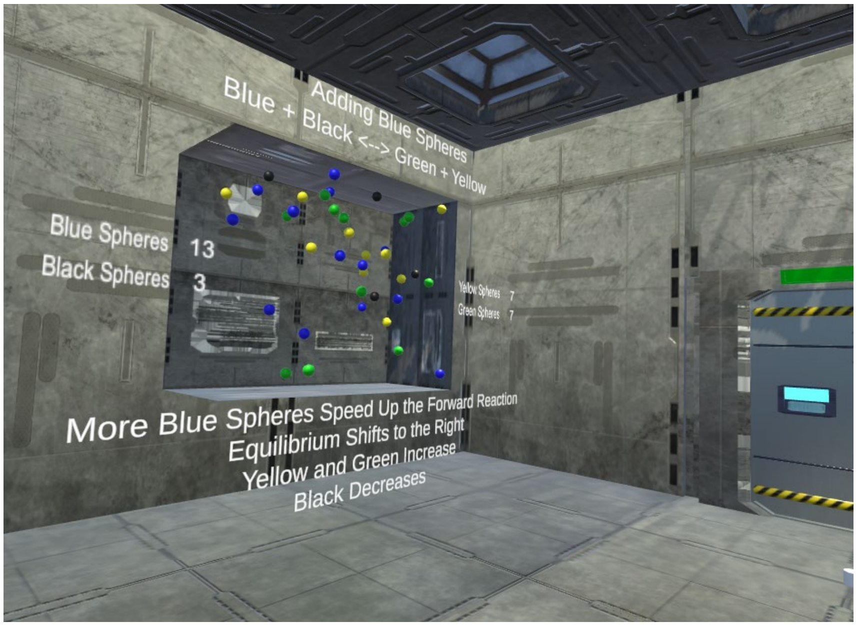

Figure 11 Debriefing Room

If players make two incorrect choices, a different door is opened, leading to a secondary pathway, which includes audio feedback (Figure 12) to enhance understanding of the challenge, followed by another reaction challenge (Figure 13). If players fail to answer the second challenge correctly, a third similar challenge is presented – following additional audio feedback. If the player solves the second challenge correctly, answer the third challenge correctly, or answer the third challenge incorrectly, the player moves through a secondary corridor and returns them to the main pathway to attempt the next challenge.

Figure 12 Feedback Room Example

Figure 13 Remedial Reaction Room

Each main challenge (and subsequent secondary challenges) introduces a different concept related to chemical equilibrium (e.g., adding/removing substances, changing temperature, changing volume/pressure). An example of the potential pathways taken through the learning environment are shown in Figure 14. Incorrect responses to challenges allow students to move along the horizontal pathways shown, while correct responses open the vertical pathways shown in the diagram. If students move horizontally, and then vertically, a connecting corridor allows the student to return to the main pathway to attempt the next challenge.

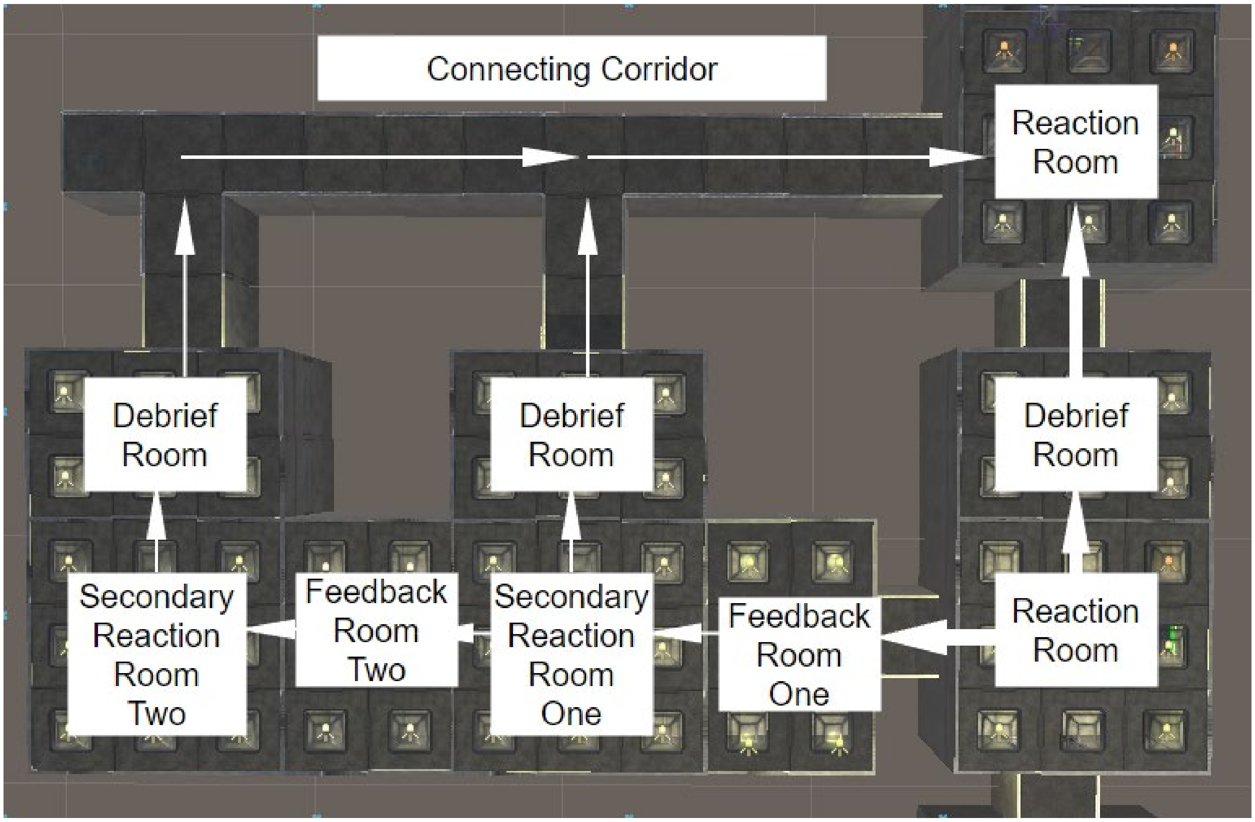

Figure 14 Diagram of Pathways Through Learning Environment

Following the fourth challenge and any secondary challenges, players have an opportunity to complete challenges for each of the key type of changes that can impact equilibrium the analogues will ultimately be replaced by systems representing actual chemical reactions (Figure 15). In order to reach and complete the final challenges, participants were required to correctly select actions (called “stresses” in the context of chemical equilibrium) that shift several different chemical reactions in a particular direction. There are a variety of different stresses that may shift equilibrium, which are assessed in the pre- and post-test content questions. Examples include adding or removing reactants, changing temperature, or changing pressure. Each type of stress was represented in both the analogues and the actual reactions demonstrated in the learning environment. Each room and each breakout challenge were designed to correlate directly to one or more of the assessment questions. After students successfully complete the final puzzle, they will escape the game. As noted in the previous sections, each room and challenge will be designed to help the student construct meaning that can be applied to the actual chemical reaction problems that will be addressed on the pre- and post-test.

Figure 15 Real Reaction Challenge

<u>ITS Design</u>

The prototype included multiple elements of intelligent tutoring systems (ITS) (Gonzales-Calero et al., 2015; Graesseret al., 2016). The structure of the environment itself was developed to provide multiple pathways for players, based on whether they solved challenges correctly or incorrectly. The shortest pathway requires students to solve 8 challenges, which would occur if a player solved each challenge correctly. If a player fails to solve any of the 8 main challenges, they proceed through a feedback room. In the feedback room, students receive audio explanations that repeat and reinforce the instructions given in the tutorial portion of the environment. After receiving this feedback concerning the topic being tested, students attempt one or two similar challenges that test the same concept. Answering one of these secondary challenges allows the player to return to the main pathway to attempt the next main challenge.

Depending on the number of challenges solved correctly or incorrectly, there are dozens of possible pathways that a player could take in order to complete the game.

In addition to the physical structure, players are given both visual and audio feedback after each challenge, whether they solve the challenge correctly or not. If a primary or secondary challenge is solved correctly, players move into a debriefing room (Figure 11) where the results of their choices are displayed in reaction chambers where the conditions have been changed based on their actions. The visual representation also includes text that explains the specific changes that have occurred. This is accompanied by audio feedback describing the changes and how they impacted the reaction.

If the player fails to solve the challenge, they move into a feedback room with they are provided with additional audio instruction on the subject being tested before attempting a secondary challenge that tests the same concept. This combination of audio, visual, and structural design provides players with dynamic and immediate feedback and support depending on their demonstrated level of understanding. In future iterations, the guidance could be provided by a tutoring “agent” in the game, like those found in commercial video games.

Description of Participants and Context

Participants

Participants were 26 volunteers from a group of 112 students in a high school chemistry course who completed the entire protocol and were included in the data analysis presented in Chapter 4. An additional eight volunteers who were not included in the analysis encountered technical issues with the application that prevented them from reaching the end of the learning environment. The sample was drawn from five sections of a course taught by a single instructor and represents approximately 10% of the college preparatory chemistry students at the high school where the study took place. The course level is described as “college preparatory” and is considered an introductory course where students establish a foundation of broad understanding of concepts in inorganic chemistry. The course is differentiated from honors and advanced placement levels of chemistry primarily in the depth of content taught, the pacing of instruction, and the rigor of expectations for mathematical foundations underlying chemical concepts.

Of the 26 students who completed the protocol, there were 5 females and 21 males. There was one student in Grade 10 and 25 students in Grade 11. The students ranged in age from 16 to 17, with 19 students being 16 years of age and the remaining 7 being 17 years of age. The proportion of male to female students is not representative of the larger population. However, the proportions by age a grade are representative of the population of College Preparatory Chemistry students within the research site. The sections selected for the study included students with diverse backgrounds in both science and mathematics, based on the variety of courses available in the science program at the school where the study will take place. Most students in the course have previously taken 1 or 2 years of biology. In addition, some students have taken a physical science course, which introduces basic chemistry and physics concepts during students’ freshman year of high school. Students who have taken physical science prior to chemistry will have been exposed to the concepts of chemical reactions and equations but have not yet studied chemical equilibrium and Le Chatelier’s Principle.

In addition to differences in prior science courses, students in these classes vary in their prior experience in mathematics. Most students in these classes have had 1 year of introductory algebra and 1 year of geometry. Students may concurrently be enrolled in Algebra II. However, a segment of this sample is also enrolled in a remedial mathematics track, involving additional instruction in algebra. While these algebraic concepts are important to many concepts in chemistry, the focus of the intervention will be on the conceptual basis for equilibrium and the ability to correctly predict shifts in concentration. Calculations of equilibrium concentrations and equilibrium constants were not required. These calculations are not required by state curriculum frameworks (Massachusetts Department of Elementary and Secondary Education, 2016).

All students in the selected course sections were provided with copies of the Informed Consent and Assent documents. The completed forms were collected by another teacher within the Chemistry Department, and students were allowed to use complete the virtual reality game regardless of their participation status. Participation status was not known to the researcher until after the completion of the project.

Setting

The demographics of the school, which is a public high school with approximately 1,500 students, is predominantly White (79%), with a minority population including Asian (12.6%), African American (2.9%), and Hispanic (3.4%). The classes selected for the study typically include approximately 2-3 students for whom English is not a native language and include 30-40% students with identified learning disabilities.

Compared to other schools in Massachusetts, the study site was within approximately the 90th percentile (89.1%) for enrollment among 11th grade students. Students at Chelmsford High School typically perform at or above state-wide averages on standardized testing for science (Massachusetts Department of Elementary and Secondary Education, 2020).

Instrumentation and Sources of Data

The study collected data on both content knowledge within the context of chemistry and user experience data for the learning environment itself. A detailed description of the instruments and data analysis are presented in the Instruments and Data Analysis sections.

Instruments

Data collected were in the form of pre- and post-test scores, as well as data collected by a post-intervention survey to assess usability. A pre-and post-test developed by the chemistry teachers within the school, based on state standards, and utilizing questions from state-mandated testing to demonstrate content knowledge (Appendix B). Questions were multiple-choice format. The Massachusetts Comprehensive Assessment Systems (MCAS) for Chemistry lists equilibrium among one of the four major categories of knowledge assessed on the test (Massachusetts Department of Elementary and Secondary Education, 2013). In addition to questions selected from released MCAS items, questions were based on the SAT Subject Matter Test in Chemistry (The College Board, 2021a), and the Advanced Placement Exam for Chemistry (The College Board, 2021b).

Each of the 10 questions included on the assessment addresses one of four factors that can shift equilibrium: 1) adding or removing a reactant, 2) adding or removing a product, 3) adding or removing heat, or 4) increasing or decreasing pressure/volume. Each of the analog and real reaction challenges in the learning environment were designed to represent one of these four changes. One analog and one real reaction challenge was created to demonstrate each of the conditions, for a total of eight main challenges. In addition to the main challenges, two secondary challenges were created for students who failed to solve the main challenge. The addition of secondary challenges resulted in a total of 24 challenges in the learning environment. In addition to the challenges replicating the concepts being tested on the assessment, the format of the challenges, in which students selected an action by placing numbered cylinders into sockets was designed to be analogous to the multiple-choice format of the post-test assessment. One key difference between the structure of the learning environment challenges and the format of the assessment was that each challenge in the learning environment was that each challenge required the student to make two correct choices to complete the challenge, or two incorrect choices to fail the challenge. This two-response model was designed to reduce the probability of a student successfully solving or failing a challenge by random chance.

Data collection

Data was collected in the form of pre- and post-test scores, using the assessment described above. In addition, qualitative data on the experience of the VR system was collected via an online post-intervention survey comprising items from the Virtual Reality Neuroscience Questionnaire (VRNQ; Kourtesis et al., 2019), the System Usability Survey (Brooke, 1996), and the User Reaction Survey (Butt et al., 2018).

The pre- and post-test assessments included questions selected from the Massachusetts Comprehensive Assessment system (MCAS) subject test in Chemistry. The MCAS was developed to establish benchmarks for “areas where students have mastered a subject, where students need more help, and where educators might need to adjust their lesson plans, materials, or approach in order for students to meet the standards” (Massachusetts Department of Elementary and Secondary Education, 2020). Questions were selected from test items released by the state and consisted of multiple-choice questions about chemical equilibrium.

All participants were administered the pre-test, prior to any instruction about chemical equilibrium. Students had prior knowledge of the fundamentals of chemical reactions, and the qualitative relationships between products and reactants. Instruction on equilibrium was exclusively delivered using the virtual reality environment. Following instruction, students were administered the post-test, consisting of the same items from the pre-test. A comparison of performance on the post-test were conducted as described below. A summary of each survey instrument is presented in Table 2.

In addition to the pre-and post-tests, students completed a survey based on the Virtual Reality Neuroscience Questionnaire to gather data about their experience in the virtual reality environment for learning. The purpose of the instrument was to create a brief survey to assess the quality the quality of VR software. The VRNQ consists of 20 questions across 4 subdomains, which measure user experience in virtual reality environments (Kourtesis et al., 2019). Criteria evaluated by the survey items include quality of graphics, sound, and hardware, navigation, naturalistic interaction with in-game objects, effective tutorials, and virtual reality induced symptoms and effects. One survey item was omitted from the VRNQ subset on Game Mechanic. The specific interaction mentioned in the item "How easy was the 2-handed interaction e.g., grab the tablet with one hand, and push the button with the other hand?" was not applicable to the game mechanics in the prototype learning environment. The survey data was analyzed for any patterns or trends that may prove of value to future iterations of the environment or future researchers.

Students also answered questions from the full System Usability Scale (Brooke, 1996), and selected items from the User Reaction Survey (Butt et al., 2018). Items from the System Usability Scale assess user perceptions of the usability of computer systems. The User Reaction Survey was developed to assess users to rate their perceptions of how a virtual reality learning tool would affect their learning. These questionnaires consist of 10 questions and 24-questions respectively use a Likert-type scale to assess user experience in various aspects of software applications and immersive virtual environments. Seven items were omitted from the User Reaction Survey because they either duplicated questions asked in previous surveys (i.e., I felt dizzy or nauseous was asked in the VRNQ), or they did not apply to the current study because they referred to specific elements of the original study that developed the instrument (i.e., "It was easy to concentrate on aseptic technique.").

Table 2 Summary of Survey Instruments

Instrument	Number of Items	Format	Source	Purpose
Pre-Test/Post-Test	10	Multiple Choice	Massachusetts Department of Education, 2020)	Assessment of Content Knowledge
Virtual Reality Neuroscience Questionnaire	19	Likert-type	Kourtesis, Collina, Doumas, & MacPherson, 2019	Measures user experience in virtual reality environments
System Usability Survey	10	Likert-type	Brooke, 1996	Assess user experience software applications
User Reaction Survey	17	Likert-type	Butt, Kardong-Edgren, & Ellertson, 2018	Assess user experience in immersive virtual environments

Data analysis

The current study collected quantitative data on student performance and user experience. The pre- and post-test data was analyzed using a paired t-test. This type of analysis is based on a null-hypothesis model. In this case, the null hypothesis was that there is no difference in scores between the pre-test and post-test and that any improvement in post-test scores can be explained by chance. If there was improvement in student performance, and the difference in the pre- and post-test scores was too large to be explained by chance, then the null hypothesis was rejected. If there is a significant difference in test scores following student interaction with the learning environment, then one interpretation of that data would be that the environment may have contributed to that difference (Creswell, 2014). In addition to the paired t-test, correlation analysis was conducted to identify any relationship between player's score in the game or time spent in the game and improvement on their performance. The analysis was conducted using RStudio for Windows, an open-source statistical software package.

For the VRNQ, participants responded to a series of questions concerning virtual reality induced symptoms and effects (VRISE), as well as user experience in a virtual reality environment (Kourtesis et al., 2019). The responses were on a 7-point Likert-type scale, in which higher scores indicated a more positive outcome. Cutoffs have been established so that a score of ≥ 25 in a subcategory and ≥ 100 overall indicates an acceptable user experience, while a subcategory score ≥ 30 and overall score ≥ 120 indicates a robust suitability for the VR environment (Somrak, Pogacnik, & Guna, 2021). Results of the user experience surveys will be used to identify features of the prototype that could be prioritized for changes in future iterations

Ethical Considerations

There were three major ethical considerations for the current study: informed consent, confidentiality, and safety. Since the participants were children, each of these considerations took on special meaning. Permission from parents or guardians, as well as school administration was required.

First, on the matter of informed consent, information concerning the study was provided to both parents and the students who will potentially participate in the study. To minimize potential bias, more information was provided to guardians than students. Guardians were informed of the objectives of the study, the activities and technologies involved in the study, and the potential risks involved in the use of virtual reality. Further, guardians were informed that all identifiable information were stripped from the data prior to analysis, and that participation was strictly voluntary. Students were provided with less details concerning the procedure of the study but were informed of the risks and voluntary nature of participation. To preserve confidentiality, participant names were removed prior to analysis. All identifiers were stripped from the data following collection, and no identifying information was stored or available to the participants or the public.

Safety considerations specific to the use of virtual reality were taken into account. First, some users experience a sense of vertigo and nausea when using immersive technologies. Technology has improved to address some of these issues, but the medical histories of participants were reviewed prior to the start of the intervention. Students who are predisposed to adverse effects of virtual reality were excluded. Students were also advised to stop using the prototype, notify their instructor, and consult with their healthcare provider in the event that they experienced any adverse effects while using the prototype. There were also considerations addressed regarding the safety of students moving in a room while wearing virtual reality headsets. The nature of the learning environment did not require students to walk in real space, but some movement was necessary, and the spatial arrangement of furniture and other students was considered when setting up the intervention. The intervention was administered to multiple students simultaneously, which further complicated safety protocols.

In order to address these safety issues, a number of steps were taken. The health issues concerning adverse effects were noted in the Consent documents and repeated by the researcher prior to any participant before they began using the prototype. The issue of movement and hazards in the room were handled by a combination of planning and features of the virtual reality technology. First, students were assigned to specific locations within the room, away from other students, and surrounded by lab benches that prevented them from crossing into another participant's space. Second, the Meta Quest 2 includes a "Guardian" feature, which includes a "Stationary Boundary" option. The Stationary Boundary creates a small area within which players can move. If a player begins to move toward or beyond the end of that area, the game stops and the view within the headset becomes translucent to show the player their position relative to the boundary. When players return to the boundary area, the game resumes.

Limitations and Future Directions

There were two aspects of the proposed system that would represent limitations and areas for future development. First, there were several game design choices for the prototype system that may have limited the effectiveness of the intervention. For example, the prototype design uses teleportation rather than full room-scale motion tracking. Full-scale tracking that would provide more realistic movement within the environment. The prototype also has only limited ability to show cause and effect of change to the reactions in real-time. Future iterations will improve on this limitation. Finally, the Intelligent Tutoring System in the prototype is a very basic version of this type of system. It consists of structural features of the virtual environment itself and audio/visual feedback. Future iterations will include more interactive and more integrated guides through the introduction of other "agents" or Non-Player Characters for the player to interact with in the environment. While these limitations do not invalidate the user experience, they do pose some risk to the immersiveness of the virtual world and the potential learning benefits of playing the game.

The second limitation to the proposed project is that each learner operated individually. Collaboration and dialogue are key elements of many revised education standards (Common Core State Standards, 2016), and the social aspect of these activities is hindered by the technology in the proposed model. More advanced technology, which has recently become available in the form of the Unity Multiplayer Networking development package, would allow for the users to communicate and interact in the virtual reality environment within the same physical space. Multiplayer serves could also be used to add players who in physically separated but virtually connected environments, offering the ability to see other users as avatars in real time. There are currently applications of this technology available in commercial virtual reality applications available for the Meta Quest 2, and a multiplayer game and training platform called Zero Latency. Future iterations of this project would be designed to take advantage of this technology, or other similar platforms.

In future iterations of the learning environment, regardless of the technology used, there are several potential directions that could enhance the success of the intervention. In addition to the pre- and -post-test data analysis, the study will analyze the responses to the VRNQ. This analysis will be conducted to examine participants' User Experience (UX) in the intervention environment. This data could be used to analyze correlation between UX and performance, which may provide valuable data for future iterations of the learning environment and can contribute to the field of design and implementation for other virtual learning systems. There are also features of the Meta Quest 2 that would allow researchers to stream the user view to another monitor, which could record the entire user experience for further analysis. Further gamification of the environment could enhance engagement, and the addition of multiplayer features would encourage collaboration and complex problem-solving aspects of the intervention.

Chapter 3 Summary

The current study involved the development of a virtual learning environment, using immersive virtual reality technology and a series of interactive scenarios that introduced students to the concepts of chemical equilibrium and Le Chatelier's Principle. The system presented students with analogs to represent the various factors that influence equilibrium, presented as puzzle-style challenges, building in complexity to scaffold learning of key concepts. A summary of the framework for the study is presented below (see Table 1 in Chapter One).

Participants were 34 high school chemistry students, primarily in the 11th grade, in a college preparatory chemistry course. The setting was a medium-sized public high school with a diverse socioeconomic and demographic population. The sections selected included between 25% and 40% students with special education services of varying levels. All sections were taught by the same instructor.

The research methodology was an educational design research model, with a pre- and post-test model for gathering performance data, and a post-intervention questionnaire to assess user experience. This design was appropriate in a setting that utilizes a convenience sample; a prototype of an intervention that can be iteratively tested, evaluated, and improved; and a research problem without a clearly-define solution. The analysis focused on the improvement of performance on the pre- and post-test, as well as user experience.

While the current study involved a diverse group of students, all research was done at a single site with a single instructor, which limits generalizability. The technology utilized is also not yet widely available and requires relatively expensive equipment, although the costs and types of equipment are changing as the technology matures. Future research may involve multi-user environments that allow for collaboration between students and instructors in the virtual environment.

CHAPTER FOUR: RESULTS

Introduction

This study designed, developed, and tested an immersive virtual reality learning environment for teaching concepts of chemical equilibrium. The testing phase of the project consisted of classroom use of the environment by students who volunteered to participate in the study. Data was collected in the form of a pre-test and post-test assessing student knowledge of chemical equilibrium concepts, as well as a survey to assess usability of the environment and the overall user experience. The data was analyzed using RStudio statistical software, and this section presents the results of these analyses.

Research Question 1: Performance Analysis

The primary research question for this study was "To what extent do high school chemistry students learn concepts of chemical equilibrium via the implementation of an immersive virtual reality environment?" In order to address this question, students completed a pre-test consisting of 10 multiple-choice questions based on standardized tests from state and federal curriculum standards. Students then completed a series of tasks using the VR learning environment. After interacting with the environment, students completed a post-test consisting of the same 10 questions on the pre-test. Students also recorded their score in the VR game and the time that they took to complete the game.

Pre-test, post-test, game scores, and times for completion were loaded into RStudio 2022.12.0 for statistical analysis. The results from the pre-test ($M = 3.27$, $SD = 1.5$) and post-test ($M = 4.77$, $SD = 2.5$) indicate that there was a significant improvement in performance on the assessment after completing the VR game, $t(25) = 4.2$, $p = 0.0003$. The effect size using Cohen's *d* was calculated to be of medium size (*d* = 0.63; 95% CI [0.30 – 0.97]; Table 3). Effect size measures the difference in means between the pre-test and post-test scores relative to the standard deviation in scores. An effect size between 0.5 and 0.8 is considered to show a medium effect. This indicates that there is a meaningful difference in the performance of students before and after using the learning environment.

Table 3 Student Performance Analysis

	Mean	**Standard Deviation**
Pre-Test Score	**3.27**	**1.5**
Post-Test Score	**4.77**	**2.5**
t	**df**	**Sig. (2-tailed)**
4.2	25	0.0003
	95% Confidence Interval of Effect Size	
Cohen's *d*	Lower	Upper
0.63	0.30	0.97

There was a slightly positive correlation between a student's score in the learning environment game and the difference between pre-test and post-test. The correlation between time to complete the game and difference in scores was slightly negative. Neither correlation was statistically significant. The correlation between improvement on the post-test score and game score was $r(24) = .16$, $p = .30$. The correlation between improvement and time spent in the game was $r(24) = -.072$, $p = .72$.

Research Question 2: Usability and User Experience Scores

The second research question investigated in the study was "What technical and conceptual challenges do students encounter while using a virtual environment for learning chemical equilibrium concepts?" In order to address this question, participants completed a post-intervention survey that consisted of a shortened version of the Virtual Reality Neuroscience Questionnaire (VRNQ; Kourtesis et al., 2019), along with adapted items from the System

Usability Scale (Brooke, 1996) and the User Reaction Survey (Butt et al., 2018). Participants responded to questions using a Likert-type scale to assess user experience in various aspects of immersive virtual environments. A summary of the response types and descriptors is presented in Table 4.

Table 4 Summary of Survey Instrument Response Types

Survey Instrument	Response Type	Low-End Descriptor	High-End Descriptor
VRNQ			
User Experience	7-point Likert-Type	1 = Extremely Low	7 = Extremely High
Game Mechanics	7-point Likert-Type	1 = Extremely Difficult	7 = Extremely Easy
In-Game Assistance	7-point Likert-Type	1 = Extremely Difficult/Unhelpful*	7 = Extremely Easy/Helpful*
VRISE	7-point Likert-Type	1 = Extremely Intense Feeling	7 = Absent
System Usability Scale	5-point Likert-Type	1 = Strongly Disagree	5 = Strongly Agree
User Reaction Survey	5-point Likert-Type	1 = Strongly Disagree	5 = Strongly Agree

* One item in the In-Game Assistance subcategory “How did you feel about the length of the tutorials?” was rated on a scale of 1 (Extremely More Time Needed) to 7 (Plenty of Time Available).

According to Kourtesis, Collina, Doumas, and MacPherson (2019), the VRNQ includes questions in four areas: User Experience, Game Mechanics, In-Game Assistance, and Virtual Reality Induced Symptoms and Effects (VRISE). Respondents answer 5 questions in each area on a 7-point Likert-type scale. The authors established both a minimum and robust score threshold for the VRNQ. A score of ≥ 25 in each area indicates an average score of 5 or better on each item in that area. Since a score of 5 is above the midpoint of 4 on each item, a total of 25 is considered the minimum cutoff to indicate that the software being tested is of adequate quality. A score of ≥ 30 indicates an average of 6 or 7 on each item and Kourtesis et al. (2019) describe as a robust indication of software quality being sufficient. There was one item in the Game Mechanics area of the survey that was not applicable to the prototype developed for this study. The item was “How easy was the two-handed interaction e.g., grab the tablet with one hand, and push the button with the other hand?” There were no two-handed interactions required in the prototype, so this item was removed from the survey. Removing the individual item reduced the minimum threshold for Game Mechanics to 20 and the overall minimum score threshold to 95.

The average score for User experience was 26.7. The average score for Game Mechanics was 23.5. The average score for In-Game Assistance was 30. The average score for VRISE was 36.4. The overall average score was 110.5. According to Kourtesis et al., a total average score of ≥ 100 is the minimum cutoff to indicate that the software was of adequate quality overall. The average scores for individual VRNQ survey items, sub-scores, and total score are shown in Table 5 below. The average scores for each survey item can be found in Appendix C

Table 5 Virtual Reality Neuroscience Questionnaire Results

Survey Area	**Average Score**
User Experience Total Score (1 = Extremely Low, 7 = Extremely High)	**26.7**
What is the level of immersion that you experienced?	5.2
What was you level of enjoyment of the VR experience?	6.1
How was the quality of the graphics?	5.0
How was the quality of the sound?	4.7
How was the quality of the VR technology overall (i.e., hardware & peripherals)?	5.7
Game Mechanics Total Score (1 = Extremely Difficult, 7 = Extremely Easy)	**23.5***
How easy was it to use the navigation system (i.e., teleportation) in the virtual environment?	5.7
How easy was it to physically move in the virtual environment?	6
How easy was it to pick up and/or place objects in the virtual environment?	5.8
How easy was it to use items in the virtual environment?	5.9
In-Game Assistance Total Score (1 = Extremely Difficult/Unhelpful, 7 = Extremely Easy/Helpful)	**30.0**
How easy was it to complete the tutorials?	6.1
How helpful were the tutorials?	6.0
How did you feel about the duration of the tutorials?	5.7
How helpful were the in-game instructions for the task you needed to perform?	6.0
How helpful were the in-game prompts e.g., arrows showing direction, or labels?	6.2
VRISE Total Score (1 = Extremely Intense Feeling, 7 = Absent)	**30.3**
Did you experience nausea?	6.3
Did you experience disorientation?	6.0
Did you experience dizziness?	6.0
Did you experience fatigue?	6.0
Did you experience instability?	6.0
Total Overall Average	**110.5**

*One question from the Game Mechanics Survey Items was not applicable to the virtual environment being studied.

The System Usability Scale (Brooke, 1996) is a 10-item survey in which respondents answer on a 5-point Likert-type scale (1= Strongly Disagree to 5 = Strongly Agree) to provide a subjective assessment of usability. Responses to each item are compiled to produce an overall usability score between 0 and 100. The User Reaction Survey (Butt et al., 2018) is a 24-item qualitative survey of user experience in which participants respond to questions on a 5-point Likert-type scale. For this survey, 17 items were selected, 9 positive items and 8 negative items. For calculating the final average score, positive questions were scored out of a possible 5 points, while negative questions were scored by subtracting the response from 5. Results of the System Usability Survey and the User Reaction Survey are included in Table 6 and Table 7 below. Individual survey items are included in Appendix B.

Table 6 System Usability Survey Results

System Usability Survey	Average Score
I think that I would like to use this system frequently.	4.2
I found the system unnecessarily complex.	3.4*
I thought the system was easy to use.	4.5
I think that I would need the support of a technical person to be able to use this system.	3*
I found the various functions in this system were well integrated.	4.3
I thought there was too much inconsistency in this system.	3.2*
I would imagine that most people would learn to use this system very quickly.	2.7
I found the system very cumbersome to use.	3.7*
I felt very confident using the system.	3
I needed to learn a lot of things before I could get going with this system.	3.7*
Total Score (Total of Survey Items x 2.5)**	**76.7**

* These items were reverse-scored. Raw score was subtracted from 5.

**Brooke notes that the System Usability Survey is intended to provide a composite score and individual item scores are not meaningful.

Table 7 User Reaction Survey Results

User Reaction Survey	
Survey Item (P = Positive; N = Negative)	**Average Score**
At times I felt totally absorbed in practicing. (P)	3.7
I did not find any challenge within this game. (N)	2.3*
Practicing this way was fun. (P)	4.5
I felt engaged in my own learning while practicing. (P)	4.5
Using this technology motivated me to keep practicing. (P)	4.5
Practicing this way was boring. (N)	3.5*
I lost track of time while practicing. (P)	3.7
Practicing this way was not engaging. (N)	3.5*
There were elements of challenge within the game. (P)	3.5
I would be more likely to practice equilibrium problems this way. (P)	4.3
I found practicing this way frustrating. (N)	3.3*
At no time was I absorbed in the game while practicing. (N)	2.8*
I found my way around the game easily. (P)	4.1
The headgear was uncomfortable. (N)	2.7*
Wearing headgear did not bother me. (P)	3.8
I did not enjoy practicing this way. (N)	3.5*
Practicing this way will help me learn chemical equilibrium. (P)	4.3
Final Average Score	**3.7**

* These items were reverse scored. Raw scores were subtracted from 5.

Research Question 3: Performance Correlation with User Experience

There was a statistically significant improvement in performance from the pre-test to the post-test (mean score improved by 15%), and the overall user experience survey results were positive. However, the third research question, "What aspects of usability and user experience correlate to student performance?" was proposed to identify any specific elements of the prototype learning environment that may have had a positive or negative correlation to student performance This question was addressed by performing a correlation analysis between usability and user experience responses and the difference in student performance. These correlations can further guide future development of the prototype and contribute to the literature by offering insights to future researchers who may be developing similar virtual reality applications.

The correlation test was run comparing items from the usability and user experience survey to the post-test scores (Table 8) and the difference in pre-test and post-test scores (Table 9). Several items showed low positive or negative correlations to either post-test scores or the difference in pre-test and post-test scores. However, only three items ("I found the system very cumbersome to use," "I found practicing this way frustrating," and "I think that I would need the support of a technical person to be able to use this system") showed a correlation with post-test scores that were significant at a level of $p < 0.05$. The correlation of the negative items was conducted on the original (non-reversed) scores. Therefore, a negative correlation indicates that the more strongly a student agreed with the negative statement, the lower their post-test score was. There was only one item ("I found practicing this way frustrating") that showed a significant correlation ($p < 0.05$) with the difference in pre-test and post-test scores. These correlations will be discussed further in Chapter 5.

Table 8 Correlation Between Performance and User Experience Survey Items

Survey Item	Correlation to Post-Test Score
I think that I would like to use this system frequently	0.37
Practicing this way will help me learn chemical equilibrium	0.31
I did not enjoy practicing this way	-0.33
Practicing this way was boring	-0.37
Wearing headgear did not bother me	-0.38
I found the system very cumbersome to use	-0.41*
I found practicing this way frustrating	-0.48*
I think that I would need the support of a technical person to be able to use this system	-0.53*

*Significant at $p < 0.05$

Table 9 Correlation Between Difference in Scores and User Experience Survey Items

Survey Item	Correlation to Score Difference
How easy was it to pick up and/or place objects in the virtual environment?	0.37
I think that I would like to use this system frequently	0.32
I think that I would need the support of a technical person to be able to use this system	-0.35
I found practicing this way frustrating	-0.42*
I needed to learn a lot of things before I could get going with this system	-0.46

* Significant at $p < 0.05$

Chapter 4 Summary

The data collected to address the research questions included a pre-test/post-test assessment and a post-test user experience survey.

The first research question was "To what extent do high school chemistry students learn concepts of chemical equilibrium via the implementation of an immersive virtual reality environment?" Findings of the pre-test and post-test assessments showed a significant difference using a paired t-test with a medium effect size. While there were small correlations between game score, time spent in the game, and improvement on the post-test assessment, these correlations were not significant.

The second research question was "What technical and conceptual challenges do students encounter while using a virtual environment for learning chemical equilibrium concepts?" Overall usability scores indicated that the VR software was of adequate quality and that the user experience was positive.

The third and final research questions was "What aspects of usability and user experience correlate to student performance?" Each of the survey items was reverse coded, (e.g., "I thought the system was easy to use." reverse coded as "I found the system very cumbersome to use."). There were no user experience survey items that showed a significant positive correlation with either post-test scores or the difference in pre-test and post-test scores. There were several negative user experience survey items that had a significant negative correlation between either post-test scores or the difference in pre-test and post-test scores. However, the corresponding reverse coded items did not show a significant positive correlation with performance. This may indicate that ease of use alone does not lead to improved performance, but difficulty in using the system is detrimental to performance. These results will be discussed in more detail in Chapter 5.

CHAPTER FIVE: DISCUSSION

Introduction

This project designed, built, tested, and evaluated the effectiveness of an immersive virtual reality learning environment for high school chemistry students to learn about concepts of chemical equilibrium. The study of chemical equilibrium using inquiry, technology, and simulations is supported in the literature (Annisa & Rohaeti, 2018; Steinkuehler & Duncan, 2008; Trey & Khan, 2009). These objectives were addressed by the primary research question:

Research Question 1: To what extent do high school chemistry students learn concepts of chemical equilibrium via the implementation of an immersive virtual reality environment?

Participant performance on a pre-test and post-test assessment was measured to address this question. The results of these assessments were presented in the Results chapter, and will be discussed in the Discussion chapter.

A secondary objective of the study was to evaluate the usability and user experience of the learning environment itself. A second and third research question were developed to address this objective:

Research Question 2: What technical and conceptual challenges do students encounter while using a virtual environment for learning chemical equilibrium concepts?

Research Question 3: What aspects of usability and user experience correlate to student performance?

These research questions will guide future development of the learning environment, as well as presenting information to other researchers and software developers working in similar fields. The results of the user experience survey were presented in the previous chapter and will be discussed in more detail below.

Research Question 1: Student Performance Key Findings

As noted in Chapter Four, the mean student score increased by an average of 1.5 points from the pre-test to the post-test. The results of the paired t-test showed a significant difference in the pre-test and post-test scores with a medium effect size. Since students received no additional teaching on the concepts being addressed in the virtual reality learning environment, any gains in understanding should be attributable to interaction with the learning environment. The application of virtual reality also made possible the visualization of reaction mechanics and equilibrium shifts that are not directly possible in the traditional laboratory setting (Twining, 2009). The data support the premise that the learning environment enables students to gain understanding of concepts of chemical equilibrium. The results of the pre-test and post-test analysis showed a significant improvement in mean scores after students completed the VR game. A more detailed examination of the results and potential weaknesses of the project is presented below.

The mean pre-test score was 3.3, indicating that students correctly answered 3 or 4 questions correctly on the pre-test assessment. It is important to note that the participants had no direct instruction on the concepts of chemical equilibrium in the current course prior to interacting with the VR game. Beyond a basic knowledge of chemical reactions as a concept and the format of chemical equations, any prior knowledge of chemical equilibrium was gained outside of the current course. With that in mind, and given the multiple-choice nature of the assessment questions, it is possible that at least some of the correct responses were the result of random guessing. It is also possible that some of the correct answers on the post-test were the result of guessing. The possibility of correct guesses is a limitation of the study in that it is not possible to evaluate how many of the correct responses on the pre-test and post-test were guesses. However, as discussed later in this chapter, there are indications that the improvements were due to knowledge gained using the VR game. The potential for improved understanding through non-traditional teaching is supported by the results of studies by Akkus et al. (2003) and Annisa and Rohaeti (2018).

Of the 26 students who completed the protocol, 4 students showed a negative difference (i.e., their post-test score was lower than their pre-test score). This decline in scores could be interpreted in one of two ways. First, students may have developed misconceptions after using the VR game that caused them to answer fewer questions correctly than they did on the pre-test. As an alternative, it is possible that the correct answers on the pre-test were the result of random guessing and those guesses were changed during the post-test, either by further random guessing, or based on information gained in the VR Game.

There was a small correlation between participants' score in the game and their improvement from pre-test to post-test, but the strength of the correlation was not significant. The score in the game was based on the number of questions the participant answered correctly. A perfect score of 25 indicated that the student answered every question correctly. Each incorrect answer reduced the final score by 1 point. However, incorrect answers triggered up to two additional questions with similar conditions. While answering questions correctly suggests that participants grasped the concepts being taught quickly, the additional exposure to questions following incorrect initial answers may have aided some students with lower scores to perform well on the post-test. This scaffolded branching structure was an intentional construct within the game.

Similarly, there was a slight negative correlation between time spent in the game and improvement from pre-test to post-test, which was not statistically significant. There could be many explanations for any connection between time spent in game and performance on the post-test. First, there was no distinction between time spent completing the tutorials and time spent in the active portion of the game. There was also no mechanism for measuring how much time was spent on each question within the game. These would be features that future iterations of the game could incorporate to gain better insight into any relationship between time spent on specific aspects of the game and performance on the assessments.

An examination of the questions answered incorrectly on the post-test may provide some insight into where the game is most effective. The questions most frequently answered correctly referred to adding heat or adding a substance to a system. The two questions most frequently answered incorrectly both referred to changing pressure. Adding heat and other substances were shown directly in the game through increasing the number of particles in the reaction chamber. This visual reference may have aided students in understanding the effects of adding and removing substances in a system. Conversely, changes in pressure were not shown directly in the game, but were shown indirectly as changes in volume. The indirect modelling of pressure may have led to misconception or difficulty in learning the concepts related to changes in pressure. Future iterations of the game will need to take this possibility into consideration.

Research Question 2: User Experience Surveys Key Findings

The second research question posed in this study was "What technical and conceptual challenges do students encounter while using a virtual environment for learning chemical equilibrium concepts?" A combination of items from the Virtual Reality Neuroscience Questionnaire (VRNQ), the System Usability Survey (SUS), and the User Reaction Survey

(URS) were administered to create an overall user experience survey. The user experience survey was administered to each student after they completed their use of the prototype and the post-test assessment.

The results of the Virtual Reality Neuroscience Questionnaire (VRNQ), the System Usability Survey (SUS) and the User Reaction Survey (URS) support the virtual reality environment developed for this study as a suitable application overall. Average scores in all three survey instruments were favorable. The favorable results are is in line with previous findings in the literature, such as Dalgarno and Lee (2010), who noted that immersive virtual reality can provide students with opportunities to learn that are superior a traditional educational setting. The VRNQ is the only survey that included guidance on minimum scores, and the results of the survey met the minimum threshold in all four categories, with two of the categories reaching the more robust parsimonious cutoff level. The overall score for the VRNQ was well above the minimum level indicating an adequate experience. The SUS and the URS do not provide minimum score thresholds, but the SUS is designed to rate a system on a scale of 0-100, and the total average score was a 76.7. The URS is rated on a scale from 1 to 5 and the total average score was 3.6. Both of these scores indicate that the overall user experience was positive. A more detailed discussion of the survey results follows below.

The VRNQ survey includes questions in four areas: User Experience, Game Mechanics, In-Game Assistance, and Virtual Reality Induced Symptoms and Effects (VRISE). The scores in all for areas were above the recommended minimum of 25 needed to support the premise that the virtual reality environment was adequately designed and built. The User Experience and Game Mechanics subscores were lower than the In-Game Assistance and VRISE subscores. The In-Game Assistance and VRISE subscores both exceeded the parsimonious threshold of 30 that indicates a more robust measure of the quality of the application. Key findings from specific survey questions will be discussed below.

The lowest average score in the User Experience area was for the item “How was the quality of the sound?” at 4.7 out of 7. The sound quality was a known issue prior to testing, as some audio files played back at low levels, making them more difficult to here. There were no environmental sounds (i.e., sound effects) and no ambient sound (i.e., soundtrack), which may have also contributed to the lower score on this item. The highest average score in the User Experience area was for the item “What was your level of enjoyment of the VR experience?” at 6.1 out of 7. All of the other survey items in this area had scores of 5 or better. The positive scores support the overall user experience as being positive.

The lowest average score in the Game Mechanics subscore items was “How easy was it to use the navigation system (e.g., teleportation) in the virtual environment?” with a score of 5.7 out of 7. While this is a positive rating, it is noteworthy for development of future iterations of the environment. One issue that arose during use of the environment was that the teleportation triggers were anchored to specific points within the game environment. The VR tracking system allowed for some degree of movement by the players, which at times resulted in players missing the trigger point. A missed trigger point in turn failed to activate other items in the game, such as doors and other teleportation points. Another issue that arose in early testing was that moving between teleportation points out of the intended sequence (e.g., skipping over a point or moving backward to a previous point) created logic loops in the game software that resulted in the game freezing. These issues resulted in some frustration on the part of players. The highest score in the Game Mechanics subscore was for the item “How easy was it to physically move in the virtual environment?” at 6.0 out of 7. Players were given freedom of movement within a limited area, and could turn, reach, crouch, and lean within those boundaries. In future iterations of the game, a larger range of motion may be incorporated, using the VR hardware's room scale tracking capabilities. All of the scores in this area were above 5.5, indicating a positive perception of the Game Mechanics.

The In-Game Assistance subscore was 30 overall, which meets the parsimonious cutoff threshold indicating a well-designed system. In-Game Assistance was one area in which the peer-testing and early pilot testing failures provided critical insights that improved the prototype environment. The highest score in this area was for the item "How helpful were the in-game prompts (e.g., arrows showing direction, or labels)?" at 6.2 out of 7. This item most likely reflected in part changes made in response to peer-testing by another chemistry teacher. Additions made to the environment after peer-testing included both visual and audio prompts to help guide players in the correct direction, as well as prompts that alerted players when they had made one correct and one incorrect choice during the game. The lowest score in this area was on the item "How did you feel about the duration of the tutorials?" with a score of 5.7 out of 7. The duration of the tutorials was designed to be approximately 10 minutes. However, the time was not restricted and players could take as much time as needed to move through the tutorials. The survey item was scaled so that a score of 5.7 falls between "Enough Time Available" and "Much Time Available." While this is a positive response, it may suggest that additional guidance would be helpful in future iterations of the tutorials. All of the other items in the In-Game Assistance subscore were 6 or greater. This data supports the assertion that the In-Game Assistance was successful.

The VRISE subscore was the highest of all subscore areas, with a total of 30.3, which is above the parsimonious cutoff threshold for a suitable VR environment. All of the items in this area had an average score of 6.0, with the exception of “Did you experience nausea?”, which had an average score of 6.3 on a scale of 1 “Extremely Intense Feeling” to 7 “Absent”. These scores indicate a very low perception of VR induced symptoms, suggesting that the game design was suitable for most users. There were some areas of concern within the game design raised by the peer-testing, including visual artifacts resulting in some objects having a jittery appearance. These were adjusted prior to pilot testing with students. The other potential area for concern was the large number of fast-moving objects in the reaction chambers. However, the favorable responses to these survey items suggest that this element of the game did not produce significant negative effects on users. Results of the VRNQ are summarized in Table 10

Table 10 Summary of Key Findings from VRNQ

Survey Area	Survey Item	Average Score	Key Takeaway
User Experience	How was the quality of the sound?	4.7	Improved sound quality and ambient sounds may enhance user experience.
	What was your level of enjoyment of the VR experience?	6.1	User experience overall was enjoyable.
Game Mechanics	How easy was it to use the navigation system (e.g., teleportation) in the virtual environment?	5.7	Improvements to teleportation system and controls should be considered.
	How easy was it to physically move in the virtual environment?	6.0	Users enjoyed freedom of movement in limited space. Room-scale movement should be considered.
In-Game Assistance	How did you feel about the duration of the tutorials?	5.7	Shorten length of tutorials or limit time available in tutorial area.
	How helpful were the in-game prompts (e.g., arrows showing direction, or labels)?	6.2	Prompts for navigation and game mechanics were effective.
VRISE	Did you experience nausea?	6.3	Environment design and hardware did not induce VR-related symptoms

The 10-item System Usability Scale (Brooke, 1996) provides a composite score between 0 and 100 to indicate a "global view of subjective assessments of usability" (p. 5). The list of 10 survey items attempt to capture a user's experience in terms of efficiency, effectiveness, and satisfaction. The overall average score on the SUS for the VR environment was 76.7, suggesting that there was a positive sense of the environment's usability. Scores on this scale ranged from a low of 35 to a high of 92.5. The interquartile range of the data was 8.75 (73.75 to 82.5), which results in a lower fence of 60.625 and an upper fence of 95.625. Applying the upper and lower fence would result in 3 outliers on the lower end of the score range (35, 55, and 55), with the 35 being an extreme outlier, falling more than 3 times the interquartile range outside of Q1. There were no outliers at the upper end of the scores.

The 17-item User Reaction Survey was developed by Butt et al. (2018) to evaluate user experiences in a virtual reality game designed for training. The survey was designed to "qualitatively measure subjects' enjoyment, engagement, and physical comfort related to the equipment, likelihood to practice" (p. 28). The overall average score on these survey items was 3.6, which suggests a positive experience, enjoyment, and likelihood to practice using this method. The survey items were divided into positive (i.e., "Practicing this way was fun") and negative (i.e., "Practicing this way was not engaging"). Responses were given on a 5-point Likert-type scale based on how strongly the participant agreed with the statement. Using this metric, a low score on a negative item is favorable.

The lowest-rated positive survey item was an average score of 3.5 on the item "There were elements of challenge within the game." How a participant defines "challenge" is of course highly subjective. However, the relatively low score on this item may indicate that future iterations of the game could incorporate more challenging elements or a variety of challenges.

The highest-rated positive survey item was an average score of 4.54 on the survey item "Practicing this way was fun." This item was followed closely by average scores of 4.5 on both "I felt engaged in my own learning while practicing" and "Using this technology motivated me to keep practicing." All three of these items indicate that the VR Game was an engaging method for learning the concepts being taught.

The two highest scores on the negative items were "At no time was I absorbed in the game while practicing" with a score of 2.2 out of 5 and "The headgear was uncomfortable" with a score of 2.27 out of 5. While both of these scores were below the midpoint of 3 on a 5-point scale, they point to potential improvements for future iterations of the game. It is worth noting that these two items may overlap to a limited degree. If a player is struggling with the comfort of the headgear, it would reduce the opportunity to being absorbed in the game. The discomfort could impact their sense of immersion in the environment (Bossard et al., 2015; Freina & Ott, 2015). There are options for improving the fit and comfort of the headgear, from better adjustment of straps to after-market headstraps with additional padding and improved weight distribution. Other possibilities for improving absorption include improvements to the audio quality and the addition of ambient sound effects or background soundtrack. A summary of the results of the User Experience Survey are presented in Table 11.

Table 11 Summary of Key Findings from User Experience Survey

Survey Item	Average Score	Key Takeaway
Positive Items		
There were elements of challenge within the game.	3.5	Improved sound quality and ambient sounds may enhance user experience.
Practicing this way was fun.	4.54	
I felt engaged in my own learning while practicing	4.5	Students felt engaged and motivated in their learning while using the learning environment.
Using this technology motivated me to keep practicing.	4.5	
Negative Items		
At no time was I absorbed in the game while practicing	2.2	Improvements to the game mechanics, comfort of the hardware, and learning environment (i.e., adding ambient sound, headphones) should be considered to increase immersion and absorption.
The headgear was uncomfortable	2.27	Consider adding after-market headstrap to improve comfort.

Research Question 3: Performance Correlation with User Experience Key Findings

The third research question for this study was "What aspects of usability and user experience correlate to student performance?" While the results of the pre-test and post-test performance showed a significant improvement, consideration was given to whether any aspects of the user experience may have had an influence on performance. For example, did students who reported a higher level of immersion perform better on the post-test? In order to address this question correlation analysis was run to look for connections between items on the user experience survey and student performance, measured either as their post-test score or the difference between the pre-test and post-test score. Any correlations identified may help to guide future iteration of the prototype and provide insight for future researchers developing immersive learning environments for education.

The three items with significant correlations to post-test scores were all negative correlations: "I found the system cumbersome to use," "I found practicing this way frustrating," and "I think that I would need the support of a technical person to be able to use this system." Each of these items could be interpreted to indicate that users struggled with either the mechanics of the game itself or the controls of the virtual reality hardware. Since these items had a negative correlation with post-test scores, it indicates that difficulty using the system itself resulted in lower scores on the post-test. These correlations suggest that future iterations of the prototype should include additional tutorials to familiarize the users with the controls and the game mechanics themselves. There may also be an indication that the game mechanics themselves were not designed intuitively. The only item that had a significant ($p < 0.05$) correlation to the difference in pre-test and post-test scores was "I found practicing this way frustrating." As noted above, this survey item might suggest that a student struggled with either the controls of the hardware, the game mechanics, or the clarity of the instructions in the system. This item was negatively correlated to the difference between pre-test and post-test scores.

All of these survey items had a negative correlation to post-test scores or difference in pre-test and post-test scores. This suggests that difficulty or frustration in using the system results in lower scores or smaller improvements on post-test assessments. It is notable that the reverse-coded items for each of the negative survey responses did not have a significant positive correlation to performance. As an example, the item "I found the system very cumbersome to use was negatively correlated to both post-test scores and pre-test to post-test improvement. This suggests that students who found the system cumbersome did not perform as well on the post-test assessment. This may indicate that the difficulty in using the system interfered with their ability to learn while using the learning environment. However, the reverse-coded item for this question was "I thought the systems was easy to use." It might be expected that, since the negative item had a negative correlation, the reverse-coded positive item would have a positive correlation to performance. This item did not have a significant correlation to student performance. The negative correlation of the negative item and the lack of a positive correlation with the positive item suggests that while difficulty using the system may reduce learning potential, ease of use alone may not be enough to produce increased learning potential.

The negative correlation could indicate a number of factors that need to be addressed in future iterations of the prototype. Since virtual reality is a highly visual medium, it can result in a high cognitive load (Souchet, 2022). Albus, Vogt, and Seufert (2021) discussed the value of adding various signaling, such as text and focus through highlighting, to reduce cognitive load and improve student recall. While text and audio guidance were provided in the prototype, the content itself was highly visual, and may have been overwhelming. In future iterations, consideration will have to be given to how much information is required to present at one time, and whether various signaling methods would help reduce cognitive load. Studies have shown that reducing cognitive load can have a positive impact on learning in virtual reality environments (Bharathan et al., 2013). Key takeaways from the correlation analysis are presented in Table 12.

Table 12 Summary of Key Findings from User Experience and Performance Correlations

<table>
<tr><th>Survey Item</th><th>Average Item Score</th><th>Correlation to Post-Test Score</th><th>Key Takeaways/Possible Improvements</th></tr>
<tr><td>I found the system cumbersome to use</td><td>2.3*</td><td>-0.41**</td><td rowspan="5">1. Students who agreed with these statements more strongly had lower scores on the post-test. Difficulty using the system negatively impacted post-test performance.
2. Consider impact of audio/visual design elements and game mechanics on cognitive load. Improve tutorials for using hardware and game mechanics, revisit environmental design to make interactions more intuitive. Improve tutorials for using hardware and game mechanics to reduce cognitive load.</td></tr>
<tr><td>I found practicing this way frustrating</td><td>1.7*</td><td>-0.48**</td></tr>
<tr><td>I think that I would need the support of a technical person to be able to use this system.</td><td>2.0*</td><td>-0.53**</td></tr>
<tr><td></td><td></td><td>Correlation to Difference in Pre-Test/Post-Test Score</td></tr>
<tr><td>I found practicing this way frustrating.</td><td>1.7*</td><td>-0.42*</td></tr>
</table>

*Raw score based on Likert-type scale from 1 – Strongly Disagree to 5 – Strongly Agree

**$p < 0.05$

Limitations and Future Considerations

This study followed an educational design research model. This design was selected and supported by other current research such as Tare, Shell, and Jackson (2020) and Morch (2020). Those studies addressed educational challenges by testing interventions and prototypes directly in the classroom setting with current students. However, the study had several key limitations that have an impact on its generalizability. First, the participants were a convenience sample of students from the researcher's classes, rather than a randomly-selected cohort. However, with the exception of the balance of males and females, the other demographics of the sample were representative of the student population of college preparatory chemistry classes at this school. The results may not be generalizable to other settings with more diverse or substantially different populations. Future studies of this application should include a larger, more representative sample. In addition, future studies could include members of different levels of high school chemistry, including honors and advanced placement classes, to see if there is any significant difference in this population.

This study was a pilot program to test a new virtual reality learning environment. As such, it followed an educational design research design, which did not include randomized selection or a control group. Future iterations of the application could follow a similar model. However, a randomized controlled trial of the technology should be conducted in the future to establish a baseline comparison for the environment's effectiveness. The design of the analogous and real chemical equations in the game was based on the topics tested in the standards-based assessment used for the pre-test and post-test. The objective was to align each challenge in the game to the learning objective reflected in the assessment (Shelton & Scoresby, 2011).

Another limitation of the study, which has been previously discussed, is that the pre-test and post-test assessments consisted of 10 multiple choice questions. This format meant that some students could have scored correct answers by guessing. As noted, there were some participants who answered questions correctly on the pre-test and answered fewer questions correctly on the post-test. This change may indicate that these students were guessing on both pre-tests and post-tests. A more open response format for the pre-test and post-test assessment could provide more insight into how much knowledge students gained by using the virtual reality game. The majority of the participants on the study were male (21) compared to female (5). This imbalance in male and female participants may limit the generalizability of the results. Further study would be needed to determine if there were significant differences in the performance between males and females, and greater participation by females would improve the generalizability.

There were several lessons learned in this pilot study that can be applied to future iterations of the VR Game, as well as providing insights for other software developers when creating immersive virtual reality learning environments. There were technical issues that will need to be addressed in any future development of the game. There were also limitations in the types of data collected for analysis that could provide more detailed feedback for future improvements. Finally, there were limitations to the features of this application as a prototype that could be researched and overcome with more time and resources designated to future iterations of the game. The features built into this version of the game did not take full advantages of the affordances of virtual reality, such as the ability to manipulate size and scale in ways that are not possible in the real world, and interactions that would not be practical in reality. Students in future versions could also collaborate and create representations of reactions to share what they have learned with peers and instructors.

First, there were a number of technical issues that arose during the peer-testing and early pilot testing of the environment. These included a range of problems with the coding of the environment itself that caused the application to crash when users executed actions in a particular unexpected order. These issues prevented some of the early participants from being able to complete the game despite several attempts to restart the application. Several changes were made to early versions of the prototype prior to further testing in an effort to eliminate these logic flaws in the coding or to prevent players from being able to perform actions outside of the parameters of the code. However, future versions of the game will need to be more flexible, in order to better accommodate unexpected actions by players. Results of the user experience survey, as noted above, indicate a negative correlation between user frustration, need for technical support, and a perception that the system was cumbersome. There are a variety of possibilities to improve on these aspects of the learning environment, from improved game mechanics, reducing cognitive load through streamlining of visual elements, to more effective tutorials on game controls, mechanics, and concepts being taught.

As noted previously, one of the aspects of the content that was most difficult to effectively display was a change in pressure. In reality, pressure and volume are correlated, and the game used volume as an analog for pressure. However, changing volume is not a direct analog, since pressure and volume have an inverse relationship. Therefore, it is possible that this created a misconception when students observed volume decreasing, which is equivalent to pressure increasing. There were also visual challenges with this aspect of the game, which caused the moving wall of the reaction chamber to be transparent when it changed the volume of the chamber. The transparency may have created additional confusion as to what was taking place in the chamber. Questions related to changes in pressure were the most frequently incorrect on the post-test assessment. Performance on questions related to pressure may have been related to the issues noted here. Future iterations of the game will need to address this issue and present changes in pressure in a way that is more directly analogous to actual reactions. Reifying the changes in pressure more accurately include designing the reaction chambers so that changes in volume are more visible. Another design option would be to include a pressure gauge to show changing pressure in real time.

Future directions for the development of the virtual reality environment include those listed above, in addition to more significant changes to the game mechanics. There are several options for changing the game that could improve the experience for players, while enhancing the learning opportunities related to the concepts being taught. As an example, rather than using simple colored spheres to represent molecules, it is possible to create models that more accurately represent the molecules involved in the reaction. Representing heat as being a part of the reaction appears to have been effective. However, adding a temperature gauge that changes in real time could enhance understanding. It is also more accurate to represent molecules changing speed when temperature is changed, which was not feasible in this pilot of the prototype. These changes may be iteratively applied to future versions of the game as the educational design research process continues through additional meso-cycles. Possible modifications to the game are discussed below.

Two key components of the game that could be modified to more directly connect player's actions to the content being taught are changes to the actions that players actually take in the game and having the game respond to player actions more directly. Each of these changes require more advanced game design that could be developed incrementally in future iterations.

In the current prototype version of the game, players indicate how they would change the system in the reaction chamber by placing cylinders into slots that represent those changes. This format was selected because in many ways it models the multiple-choice format of the pre-test and post-test assessments. However, in terms of the action performed by the students, there is no difference between selecting a change in pressure and adding a reactant. Each choice is made by placing a cylinder into a slot. In future iterations of the game, players could actually perform the action that they select (i.e., add more heat) by physically placing more heat spheres into the chamber or change the pressure in the chamber by operating a dial or lever. Introducing opportunities to have direct influence on the system would provide a direct connection between the action they perform and its effect on the system. Direct interaction would remove one layer of abstraction from the learning process and allow the player to more directly influence the game. Ahn et al. (2015) found that, even in an immersive environment, direct interaction and participation had a longer-lasting effect than simply observing an action in virtual reality.

Another aspect of the game that could be changed to make the experience more direct is to have the system respond to changes immediately following the action selected by the player. In the current version, players make their selections and then move to another room that shows the results of their selections. The design was chosen as part of the decision to require two correct answers to solve a challenge, while also providing students with the opportunity to change their selections. If the system responded immediately, students would have the opportunity to find the correct responses by process of elimination alone. However, this before/after design creates a delay between making the choice and seeing the effects of their selections. Seeing the results of their actions in this format may limit a player's ability to connect cause and effect. If the system immediately and accurately changed in response to player correct and incorrect actions, it would reinforce the connection and potentially improve understanding. If this modification could be made in conjunction with the change noted above to allow players to directly model the changes they want to make, it could significantly improve student understanding.

Future iterations of the game will also be guided by prior and emerging research (Baker & Delacruz, 2016; Bartle, 2004; Shelton, Satwicz, & Caswell, 2011; Zikas et al., 2020). New possibilities are emerging for opportunities to collaborate in virtual reality. Multiplayer collaboration could be incorporated into the educational game using the Unity development environment, allowing students to work together to solve problems.

Conclusion

The current project was carried out to design, build, and test an immersive virtual reality learning environment to teach high school students' concepts of chemical equilibrium.

The research was designed to address three research questions as follows:

Research Question 1: To what extent do high school chemistry students learn concepts of chemical equilibrium via the implementation of an immersive virtual reality environment?

Research Question 2: What technical and conceptual challenges do students encounter while using a virtual environment for learning chemical equilibrium concepts?

Research Question 3: What aspects of usability and user experience correlate to student performance?

The first objective of the study was achieved. A prototype virtual reality game was designed and built using off-the-shelf development software and standalone virtual reality hardware. Following testing by a high school chemistry teacher familiar with both virtual reality gaming and the concepts being taught, a prototype was pilot-tested using volunteer students in a college preparatory chemistry course. A total of 26 participants completed a pre-test assessment, played the virtual reality game, and then completed a post-test and user experience survey. Pre-test and post-test scores were compared to measure learning gains made by playing the virtual reality game. The results of the pre-test/post-test showed a significant improvement in the average scores on the assessments. The results of the usability survey indicated that participants found the experience of using the virtual reality application

The overall results address the primary research questions, supporting the concept that students can learn about concepts of chemical equilibrium by interacting with an immersive virtual reality learning environment. The overall improvement in performance, was significant, but there is room for improvement in future iterations of the game. The overall results of the usability surveys and issues with early pilot testing indicate several areas in which users faced challenges when using the system. These results will help to guide improvements and additions to future versions of the application as the educational design research continues.

This project has demonstrated that students with limited prior knowledge can learn about complex and challenging chemistry concepts through the introduction of a prototype virtual reality game. The success of the project suggests that there is an opportunity for further investigation into improving student understanding of this project, and students may benefit from similar learning environments to teach other complex science concepts. This study contributes to the field by demonstrating that immersive virtual reality learning environments have the potential to improve teaching and learning for high school chemistry students.

References

Ahn, S. J., Fox, J., Dale, K. R., & Avant, J. A. (2015). Framing virtual experiences: Effects on environmental efficacy and behavior over time. Communication Research, 42(6), 839-863.

Akkus, H., Kadayifci, H., Atasoy, B., & Geban, O. (2003). Effectiveness of instruction based on the constructivist approach on understanding chemical equilibrium concepts. *Research in Science & Technological Education*, *21*(2), 209-227, doi: 10.1080/0263514032000127248

Albus, P., Vogt, A., & Seufert, T. (2021). Signaling in virtual reality influences learning outcome and cognitive load. *Computers & Education, 166*, 1-16.

Annisa, D. & Rohaeti, E. (2018). The effect of inquiry-based learning on students' understanding of the chemical equilibrium concept. *AIP Conference Proceedings*, 2021, 80004.

Baker, E. L., & Delacruz, G. C. (2016). A framework to create effective learning games and simulations. In H.F. O'Neil, E. L. Baker, & R. S. Perez (Eds.). *Using games and simulations for teaching and assessment: Key issues*. New York: Routledge.

Banerjee, A. (1995) Teaching chemical equilibrium and thermodynamics in undergraduate general chemistry classes. *Journal of Chemical Education*, *72*, 879-881.

Bartle, R. A. (2004). *Designing virtual worlds*. Berkeley, CA: Pearson.

Basur, M., & Durmus, S. (2010). The effectiveness of computer supported versus real laboratory inquiry learning environments on the understanding of direct current electricity among pre-service elementary school teachers. *Eurasia Journal of Mathematics, Science & Technology Education*, *6*, 47-61.

Becker, B. (2013) Learning Analytics: Insights into the Natural Learning Behavior of Our Students. *Behavioral & Social Sciences Librarian*, *32*(1), 63-67.

Bharathan, R., Vali, S., Setchell, T., Miskry, T., Darzi, A., & Aggarwal, R. (2013). Psychomotor skills and cognitive load training on a virtual reality laparoscopic simulator for tubal surgery is effective. *European Journal of Obstetrics, Gynecology, and Reproductive Biology,169*(2), 347–352. https://doi.org/10.1016/j.ejogrb.2013.03.017

Blascovich, J., & Bailenson, J. (2011). *Infinite reality: The hidden blueprint of our virtual lives*. New York: HarperCollins.

Bombari, D., Schmid Mast, M., Canadas, E., & Bachmann, M. (2015). Studying social interactions through immersive virtual environment technology: virtues, pitfalls, and future challenges. *Frontiers in Psychology*, *6*, 869. http://doi.org/10.3389/fpsyg.2015.00869

Bossard, C., Kermarrec, G., Buche, C., & Tisseau, J. (2008). Transfer of learning in virtual environments: A new challenge? *Virtual Reality*, *12*, 151-161.

Brom, C., Pruess, M., & Klement, D. (2011). Are Educational Computer Micro-Games Engaging and Effective for Knowledge Acquisition at High-Schools? A Quasi-Experimental Study. *Computers & Education*, *57*, 1971-1988.

Brooke, J. (1996). SUS: A “quick and dirty” usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (Eds.), *Usability Evaluation in Industry*. London: Taylor and Francis.

Bursztyn, N., Pederson, J., Shelton, B. E., Walker, A., & Campbell, T. (2015). Utilizing geo-referenced mobile game technology for universally accessible virtual geology field trips. *International Journal of Education in Mathematics, Science and Technology*, *3*(2), 93-100.

Bursztyn, N., Walker, A., Shelton, B. E. & Pederson, J. (2017). Assessment of student learning using virtual Grand Canyon field tips for mobile smart-devices. *Geosphere*, *13*(2), 260-268. doi:10.1130/GES01404.1

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. https://doi.org/10.1016/j.ecns.2017.09.010

Callaghan, M., McCusker, K., Losada, J. L., Harkin, J. and Wilson, S. (2009). Integrating virtual worlds & virtual learning environments for online education, *2009 International IEEE Consumer Electronics Society's Games Innovations Conference*, London, 2009, pp. 54-63. doi: 10.1109/ICEGIC.2009.5293576

Campbell, D. T. and Stanley, J. C. (1966). *Experimental and Quasi-experimental designs for research*. Retrieved from https://www.sfu.ca/~palys/Campbell&Stanley-1959-Exptl&QuasiExptlDesignsForResearch.pdf

Chani, F., Ngcoza, K. M., Chikunda, C., & Sewry, J. (2018). Exploring the mediation of learning of chemical equilibrium to high-achieving students in a selected secondary school in Namibia. *African Journal of Research in Mathematics, Science, and Technology Education*, *22*(3), 287-296, doi: 10.1080/18117295.2018.1528-31

Chen, C.-M. (2009). Personalized e-learning systems with self-regulated learning assisted mechanisms for promoting learning performance. *Expert System with Applications*, *36*, 8816-8829.

Cheung, D. (2009). The adverse effects of Le Chatelier's Principle on teacher understanding of chemical equilibrium. *Journal of Chemical Education*, *86*(4), 514-518.

Common Core State Standards (2016). *Key shifts in English Language Arts and Mathematics*. Retrieved from http://www.corestandards.org/other-resources

Creswell, J. W. (2014) *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: Sage.

Dalgarno, B., & Lee, M.J.W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, *41*(1), 10-32.

Dass, S., Dabbagh, N., & Clark, K. (2011). Using Virtual Worlds: What the Research Says. *Quarterly Review of Distance Education*, *12*(2), 95-111.

Davenport, J. L., Leinhardt, G., Greeno, J., Koedinger, K., Klahr, D., Karabinos, M., & Yaron, D. J. (2014). Evidence-based approaches to improving chemical equilibrium instruction. *Journal of Chemical Education*, *91*, 1517-1525.

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

Dickey, M. D. (2011). Murder on Grimm Isle: The impact of game narrative design in an educational game-based learning environment. *British Journal of Educational Technology*, *42*, 456-469.

Facebook Technologies (2018). Retrieved from https://support.oculus.com/170128916778795/

Facebook Technologies (2019). *Overview*. Retrieved from https://www.oculus.com/quest/.

Fakinlede, I., Kumar, V., Wen, D., & Kinshuk (2013). Conversational Forensics: Building Conversational Pedagogical Agents with Attitude. In Kinshuk & Iyer, S. (Eds.) *2013 IEEE Fifth International Conference on Technology for Education, Kharagpur, India*, (pp. 65-68), IEEE.

Finkelstein, N. D., Adams, W. K., Keller, C.J., Kohl, P. B., Perkins, K. K., Podolefsky, N. S., & Reid, S. (2005) When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment. *Physics Education Research, 1*, 1- 8.

Fogarty, I., Geelan, D., & Mukherjee, M. (2012). Does teaching sequence matter when teaching high school chemistry with visualizations? *Teaching Science*, *58*(3), 19-23.

Freeman, A., Adams Becker, S., Cummins, M., Davis, A., & Hall Geisinger, C. (2017). *NMC/CoSN Horizon Report 2017 K-12 Edition.* Austin, Texas: The New Media Consortium.

Freina, L. & Ott, M. (2015). A Literature Review on Immersive Virtual Reality in Education: State of The Art and Perspectives. Proceedings from eLSE 2015: *The 11th International Scientific Conference "eLearning and Software for Education."* Bucharest, Romania: National Defence University Publishing House

Gagne, R. M. (1965). *Conditions of learning*. New York: Holt, Rinehart, and Winston.

Ghirardi, M., Marchetti, F., Pettinari, C., Regis, A., & Roletto, E. (2014). A teaching sequence for learning the concept of chemical equilibrium in secondary school education. *Journal of Chemical Education*, *91*, 59-65.

Girvan, C., & Savage, T. (2019). Virtual Worlds: A New Environment for Contructionist Learning. *Computers in Human Behavior*, *99*, 396-414.

Glowa, L., & Goodell, J. (2016). *Student-centered learning: Functional requirements for integrated systems to optimize learning*. Vienna, VA: International Association for K-12 Online Learning.

Goldstone, W. (2011). *Unity 3.x game development essentials: Game development with C# and Javascript*. Birmingham, UK: PACKT Publishing.

Gonzalez-Calero, J. A., Arnau, D., Puig, L., & Arevalillo-Herraez, M. (2015). Intensive scaffolding in an intelligent tutoring system for the learning of algebraic word problem solving. *British Journal of Educational Technology*, *46*(6), 1189-1200.

Graesser, A. C., Hu, X., Nye, B. D., & Sottilare, R. A. (2016). Intelligent tutoring systems, serious games, and the generalized intelligent framework for tutoring (GIFT). In H. F.

O'Neil, E. L. Baker, & R. S. Perez (Eds.), *Using games and simulations for teaching and assessment* (pp.99-142). New York: Routledge.

Hanson, K., & Shelton, B. E. (2008). Design and Development of Virtual Reality: Analysis of Challenges Faced by Educators. *Educational Technology & Society*, *11*(1), 118-131.

Hew, K. F. & Cheung, W. S. (2010). Use of three-dimensional (3-D) immersive virtual worlds in K-12 and higher education settings: A review of the research. *British Journal of Educational Technology*, *41*(1), 33-55. doi:10.1111/j.1467-8535.2008.00900.x

Hwang, G.-J. (2003). A conceptual map model for developing intelligent tutoring systems. *Computers & Education*, *40*(3), 217-235.

Johnson, L., Adams Becker, S., Estrada, V., Freeman, A. (2014). *NMC Horizon Report: 2014 Higher Education Edition*. Austin, Texas: The New Media Consortium.

Kelley, A. E. (2009). When is design research appropriate? In J. van den Akker, B. Bannan, A. E. Kelly, N. Nieveen, & T. Plomp (Eds.), *An Introduction to Educational Design Research: Proceedings of the seminar conducted at the East China Normal University, Shanghai* (pp. 9-35). Enschede, The Netherlands: SLO - The Netherlands Institute for Curriculum Development.

Kinshuk (2016). *Designing adaptive and personalized learning environments*. New York: Routledge.

Kirkley, S. E., & Kirkley, J. R. (2007). Creating next generation blended learning environments using mixed reality, video games, and simulations. *TechTrends*, *49*(3), 42-53.

Kirkwood, A., & Price, L. (2013). Examining some assumptions and limitations of research on the effects of emerging technologies for teaching and learning in higher education. British Journal of Educational Technology, 44(4), 536–543. https://doi-org.libproxy.boisestate.edu/10.1111/bjet.12049

Kourtesis, P., Collina, S., Doumas, L. A. A., & MacPherson, S. E. (2019) Validation of the Virtual Reality Neuroscience Questionnaire: Maximum Duration of Immersive Virtual Reality Sessions Without the Presence of Pertinent Adverse Symptomatology. *Frontiers in Human Neuroscience*, *13*,417. doi: 10.3389/fnhum.2019.00417

Kynigos, C. & Futschek, G. (2015). Re-Situating Constructionism. *Constructivist Foundations*, *10*(3), 281-284.

Lintean, M., Rus, V., & Azevedo, R. (2011). Automatic detection of student mental models based on natural language student input during metacognitive skill training. *International Journal of Artificial Intelligence in Education*, *21*, 169-190.

Lucanus, C. (2011). A case for de-emphasizing Le Chatelier's Principle in high school chemistry courses. *Teaching Science*, *57*(4), 51-52.

Maia, P. F., & Justi, R. (2009). Learning of chemical equilibrium through modelling-based teaching. *International Journal of Science Education*, *31*(5), 603-630, doi: 10.1080/09500690802538045

Massachusetts Department of Elementary and Secondary Education (2013). *MCAS 2013 High School Chemistry Released Items*. Malden, MA: Author.

Massachusetts Department of Elementary and Secondary Education (2016). *2016 Massachusetts Science and Technology/Engineering Curriculum Framework*. Malden, MA: Author.

Massachusetts Department of Elementary and Secondary Education (2020). *MCAS High School Science and Technology/Engineering Test Results*. Retrieved from http://profiles.doe.mass.edu/mcas/science_participation.aspx?linkid=28&orgcode=00560505&fycode=2019&orgtypecode=6&

McKenney, S., & Reeves, T. C. (2012). *Conducting educational design research*. New York: Routledge.

Mei, H. H., & Sheng, L. S. (2011). Applying situated learning in a virtual reality system to enhance learning motivation. *International Journal of Information and Education Technology*, *1*, 298-302/

Mercan, Ş. & Becerikli, Y. (2020). Agile Methods in Game Programming based on Scrum . Sakarya University Journal of Science , 24 (5) , 882-891 . DOI: 10.16984/saufenbilder.658752

Mørch, A. I. (2020). Two 3D virtual worlds as domain-oriented design environments: closing the educational gap with the action-breakdown-repair model. International Journal of Information & Learning Technology, 37(5), 295–307. https://doi-org.libproxy.boisestate.edu/10.1108/IJILT-03-2020-0029

National Research Council. (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Washington, DC: The National Academies Press.

Nelson, B. C., & Erlandson, B. E. (2012). *Design for learning in virtual worlds*. New York: Routledge.

NGSS Lead States. 2013. *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.

Noss, R., & Clayson, J. (2015). Reconstructing Constructionism. Constructivist Foundations, 10(3), 285–288.

Oncu, S. and Cakir, H., (2011). Research in online learning environments: Priorities and methodologies. *Computers & Education*, *57*, 1098-1108. [Article can be accessed from Boise State library]

Padiotis, I., & Mikropoulos, T. A. (2010). Using SOLO to evaluate an educational virtual environment in a technology education setting. *Educational Technology & Society*, *13*(3), 233-245.

Papert, S., & Harel, I. (1991). *Constructionism*. Newark, NJ: Ablex Publishing Corporation.

Passig, D. (2015). Revisiting the Flynn Effect through 3D immersive virtual reality (IVR). *Computers & Education*, *88*, 327-342.

Perez, R. S., Skineer, A., & Chatelier, P. (2016). Lessons learned from intelligent tutoring research for simulation. In H. F. O'Neil, E. L. Baker, & R. S. Perez (Eds.), *Using games and simulations for teaching and assessment* (pp.99-142). New York: Routledge.

Plomp, T. (2009). Educational Design Research: An Introduction. In J. van den Akker, B. Bannan, A. E. Kelly, N. Nieveen, & T. Plomp (Eds.), *An Introduction to Educational Design Research: Proceedings of the seminar conducted at the East China Normal*

University, Shanghai (pp. 9-35). Enschede, The Netherlands: SLO - The Netherlands Institute for Curriculum Development.

Primastuti, M., & Atun, S. (2018). Analysis of students' science literacy concerning chemical equilibrium. *AIP Conference Proceedings, 2021*, 80006.

Reeves, T. C., & Amiel, T. (2008). Design-Based Research and Educational Technology: Rethinking Technology and the Research Agenda. *Educational Technology & Society*, *11*(4), 29-40.

Richards, D., & Taylor, M. (2015). A comparison of learning gains when using a 2D simulation tool versus a 3D virtual world: An experiment to find the right representation involving the Marginal Value Theorem. *Computers & Education*, *86*, 157-171.

Rubio-Tamayo, J. L., Barrio, M. G., & García, F. G. (2017). Immersive Environments and Virtual Reality: Systematic Review and Advances in Communication, Interaction and Simulation. *Multimodal Technologies and Interactions, 1*(4), 21. doi:10.3390/mti1040021

Rus, V., & Stefanescu, D. (2016). Non-intrusive assessment of learner's prior knowledge in dialogue-based intelligent tutoring systems. *Smart Learning Environments*, *3*(2). Retrieved from http://slejournal.springeropen.com/articles/10.1186/s40561-016-0025-3

Rus, V., D'Mello, S., Hu, X., & Graesser, A. C. (2013). Recent advances in conversational intelligent tutoring systems. *AI Magazine*, *3*, 42-54.

Schiaffino, S., Garcia, P., & Amandi, A. (2008). eTeacher: Providing personalized assistance to e-learning students. *Computers & Education*, *51*, 1744-1754.

Scoresby, J. & Shelton, B. E. (2011). Visual perspectives within educational computer games: Effects on presence and flow within virtual learning environments. *Instructional Science*, *39*(3), 227-254.

Shadish, W. R., Cook, T. D., and Campbell, D. T. (2002). *Quasi-Experimental designs for generalized causal inference*. Retrieved from https://www.alnap.org/system/files/content/resource/files/main/147.pdf

Shelton, B. E., & Hedley, N. R. (2003). Exploring a cognitive basis for learning spatial relationships with augmented reality. *Technology, Instruction, Cognition & Learning*, *1*(4). 323-357.

Shelton, B. E., Satwicz, T. & Caswell, T. (2011). Historical perspectives on games and education from the learning sciences. *International Journal of Game-Based Learning, 1*(3), 84-96.

Shelton, B. E., & Scoresby, J. (2011). Aligning Game Activity with Educational Goals: Following a Constrained Design Approach to Instructional Computer Games. *Educational Technology Research & Development*, *59*(1), 113-138.

Shute, V., & Ventura, M. (2009). *Stealth assessment: Measuring and supporting learning in video games*. Cambridge, MA: MIT Press.

Smaldone, R. A., Thompson, C. M., Evans, M., & Voit, W. (2017). Teaching science through video games. *Nature Chemistry*, *9*(2), 97-102.

Smith, T. R., & Gilbert, J. E. (2019). Did Math Make Me Move? The Design and Initial Evaluation of a Culturally Appropriate Gestural Educational Technology (Research). *Proceedings of the ASEE Annual Conference & Exposition*, 6770–6787.

Somrak, A., Pogačnik, M., Guna, J., & Distante, C. (2021). Suitability and Comparison of Questionnaires Assessing Virtual Reality-Induced Symptoms and Effects and User Experience in Virtual Environments. Sensors (14248220), 21(4), 1185. https://doi-org.libproxy.boisestate.edu/10.3390/s21041185

Souchet, A. D. (2022). Measuring Visual Fatigue and Cognitive Load via Eye Tracking while Learning with Virtual Reality Head-Mounted Displays: A Review. *International Journal of Human-Computer Interaction, 38*(9), 801.

Southgate, E., & Smith, S. (2017). Designing and conducting research using immersive technologies in schools: Seven observations, presented at *IEEE Xplore Virtual Reality (VR),* Los Angeles, 2017. Los Angeles, CA: IEEE.

Southgate, E., Smith, S., & Scevak, J. (2017). Asking ethical questions in research using immersive virtual and augmented reality technologies with children and youth, presented at *IEEE Xplore Virtual Reality (VR).* Los Angeles, 2017. Los Angeles, CA: IEEE.

Ştefan, L. (2012). Immersive Collaborative Environments for Teaching and Learning Traditional Design. *Procedia - Social and Behavioral Sciences*, *51*, 1056-1060. https://doi.org/10.1016/j.sbspro.2012.08.287

Steiner, P. M., Wroblewski, A., & Cook, T. D. (2009). Randomized experiments and quasi-experimental designs in educational research. In K. Ryan and J. B. Cousins (eds.), *The Sage International Handbook of Educational Evaluation* (pp.75-96). Thousand Oaks, Calif.: Sage Publications. Retrieved from http://www.ipr.northwestern.edu/workshops/past-workshops/quasi-experimental-design-and-analysis-in-education/recent-articles/docs/QE2010-steiner_wroblews_cook_handbook09.pdf

Steinkuehler, C., & Duncan, S. (2008). Scientific habits of mind in virtual worlds. Journal of *Science Education and Technology*, *17*, 530-543.

Tare, M., Shell, A. R., & Jackson, S. R. (2020). Student engagement with evidence-based supports for literacy on a digital platform. *Journal of Research on Technology in Education*, *1*. https://doi-org.libproxy.boisestate.edu/10.1080/15391523.2020.1821412

The College Board (2021a). *Chemistry Subject Test*. Retrieved from https://collegereadiness.collegeboard.org/sat-subject-tests/subjects/science/chemistry

The College Board (2021b). *AP Chemistry: Essential resources for AP Chemistry*. Retrieved from https://apcentral.collegeboard.org/

Trey, L., & Khan, S. (2008). How science students can learn about unobservable phenomena using computer-based analogies. *Computers & Education*, *51*, 519-529.

Turucz, E., Truta, S. T., Paşca, M. D., Jerzicska, E., Boeriu, C. M., Szederjesi, J., & Azamfirei, L. (2021). Educational design research, an innovative method to investigate virtual reality applications in disaster management training programs. Acta Marisiensis. Seria Medica, 67(1), 47–54. https://doi-org.libproxy.boisestate.edu/10.2478/amma-2021-0003Twining, P. (2009). Exploring the educational potential of virtual worlds. *British Journal of Educational Technology*, *40*(3), 496-514.

Tyson, L., & Treagust, D. F. (1999). The complexity of teaching and learning chemical equilibrium. *Journal of Chemical Education*, *76*(4), 554-558.

Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds). (2006). *Educational design research*. London: Routledge.

Van Eck, R. (2007). Six ideas in search of a discipline. In B. E. Shelton & D. A. Wiley (Eds.), *Educational design & use of computer simulation games* (pp. 31-60). Rotterdam, The Netherlands: Sense Publishers.

Van Lehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197-221.

Virtual reality. (2018). In *Oxford Living Dictionary*. Retrieved from https://en.oxforddictionaries.com/definition/virtual_reality

Wang, D., Han, H., Zhan, Z., Xu, J., Liu, Q., & Ren, G. (2015). A problem solving oriented intelligent tutoring system to improve students' acquisition of basic computer skills. *Computers & Education*, *81*, 102-112.

Waqar, U. (2020). Agile Scrum Crash Course: *A Guide to Agile Project Management and Scrum Master Certification PSM 1*. KDP Print.

Weir, L. (2008). Get a life: Students collaborate in simulated roles. *Edutopia*. Retrieved from http://www.edutopia.org/second-life-virtual-reality-collaboration

Wilensky, U., & Reisman, K. (2006). Thinking Like a Wolf, a Sheep, or a Firefly: Learning Biology Through Constructing and Testing Computational Theories—An Embodied Modeling Approach. *Cognition & Instruction*, *24*(2), 171-209.

Willis, J. (2001). Foundational assumptions for information technology and teacher education. *Contemporary Issues in Technology and Teacher Education*, *1*(3). Retrieved from https://www.citejournal.org/volume-1/issue-3-01/editorial/new-page-3-2

Zhu, Z., Yu, M., & Riezobos, P. (2016). A research framework of smart education. *Smart Learning Environments*, *3*(1). Retrieved from http://slejournal.springeropen.com/articles/10.1186/s40561-016-0026-2

Zikas, P., Papagiannakis, G., Lydatakis, N., Kateros, S., Ntoa, S., Adami, I., & Stephanidis, C. (2020). Immersive visual scripting based on VR software design patterns for experiential training. Visual Computer, 36(10–12), 1965–1977. https://doi-org.libproxy.boisestate.edu/10.1007/s00371-020-01919-0

APPENDIX A

IRB Application

BOISE STATE UNIVERSITY

DIVISION OF RESEARCH & ECONOMIC DEVELOPMENT
Office of Research Compliance | Institutional Review Board
(phone) 208.426.5401| MS 1138
humansubjects@boisestate.edu

EXEMPT PROTOCOL APPLICATION

INSTRUCTIONS

- The application must be typed. **Handwritten applications will not be accepted.**
- **All Questions MUST be answered.**
- **SUBMIT COMPLETED APPLICATION TO: HUMANSUBJECTS@BOISESTATE.EDU**

SECTION A: General Information

1. This project is: X Social Behavioral ☐ BioMedical
2. Project Title: **A New Reality in Teaching Chemistry: Developing a Virtual Reality Intervention for Teaching and Learning Concepts of Chemical Equilibrium**
3. Anticipated Start Date: **September 1, 2022** Anticipated End Date: **October 31, 2022**
4. **PRINCIPAL INVESTIGATOR (PI)** (Refer to the IRB PI Eligibility requirements. IRB staff will confirm your eligibility. **Graduate thesis or dissertation students MUST list eligible faculty as PI – student may be a Co-Principal Investigator.**

Name: **Dr. Brett Shelton**

Title: X Full Professor ☐ Associate Professor ☐ Assistant Professor
☐ External Investigator ☐ *Other:
*(This category is allowed under ***special circumstances only***)

☐ If you DO NOT fall into one of the above titles, you may be a Co-Principal Investigator (Co-PI) with a qualified PI.

Department: **Educational Technology** Phone: **208.426.3391**

E-mail: **bshelton@boisestate.edu**

Roles and responsibilities in this study: Dissertation Advisor

CITI Training Completed: X Social & Behavioral Researchers ☐ Biomedical Researchers

5. **CO-INVESTIGATOR** (IRB staff will confirm your title with the directory.)
Name: **Jason Ward**

☐	Full Professor	☐	Associate Professor	☐	Assistant Professor
☐	Adjunct Faculty	☐	Lecturer	☐	Undergraduate Student
☐	Staff	**X**	Graduate Student	☐ Thesis	X Dissertation
☐	Other:				

Department: **Educational Technology** Phone: **781-264-2607**
E-mail: **jasonrward@gmail.com**
Roles and responsibilities in this study: Investigator

CITI Training Completed: X Social & Behavioral Researchers ☐ Biomedical Researchers

6. Do you have additional research personnel (Co-Investigators, key personnel, student research assistants, etc.)?
X NO
☐ YES
To list additional investigators and/or key personnel, complete and attach an **ADDITIONAL PERSONNEL** form.

SECTION B: Financial Conflict of Interest Disclosure

Conflicts of interest must be disclosed in accordance with the Boise State Conflict of Interest and Commitment Policy #1110.

1. Do any investigators (PI, Co-Investigator) or research team members (key personnel) have any relationship or equity interest with any institutions or sponsors related to this research that might present or appear to present a conflict of interest (COI) with regard to the outcome of the research?
X NO POTENTIAL CONFLICTS EXIST
☐ YES:
2. Name of the person(s) with the potential COI: ______
☐ This potential conflict has been disclosed to the Boise State Conflict of Interest Office.
☐ This conflict has not been disclosed to the Boise State COI Office.

Note: If a significant conflict of interest exists, you must also attach the Boise State COI Committee approved management plan. If you have questions about conflicts of interest, contact the Boise State Conflict of Interest Officer at (208) 426-1252.

3. Is this research supported in whole or in part by a grant or contract?
X NO
☐ YES:
Sponsor Name: ______
PI on Grant: ______

Grant Title/Contract: ____________________
Project Period: From: ___ To: ______
☐ Grant Project Summary Attached
OSP Proposal Number (if known): __________________

Is this research funded internally by Boise State University?
X NO ☐ YES:

SECTION C:	YES	NO
Will the data be recorded by the investigator in such a manner that the identity of the subjects can be readily ascertained **OR** be potentially damaging to a participant's financial standing, employability or reputation?	☐	**X**
Will your research participants include prisoners, cognitively, economically, or educationally impaired participants?	☐	**X**
Will the information be obtained in such a manner that the identity of the participant can be readily ascertained, directly or through identifiers, linked to the subjects? (Exempt Category 2 or 3-requiring Limited IRB Review)	☐	**X**
Does the research involve **federal** department or agency heads for the purpose of assessing or changing public benefit or service programs? (Exempt Category 5)	☐	**X**
Does the research involve the *storage or maintenance* of identifiable private information or bio-specimens? (Exempt Category 7 – requires Limited IRB Review)	☐	**X**
Does the research involve *using* identifiable private information or identifiable bio-specimens? (Exempt Category 8 – requires Limited IRB Review)	☐	**X**

- ***If you answered YES to any of these questions, your application does NOT qualify for exempt review. STOP COMPLETING THIS FORM and complete the Expedited or Full Board Protocol Application for IRB review.***

SECTION D: Exempt Research Category

Exempt Categories 5, 7, and 8 require Expedited Application. Study must fit *exactly* into one of these categories.

Indicate the applicable Exempt Category (1-4 or 6):

X	**1.**	**Research, conducted in established or commonly accepted educational settings that specifically involves normal educational practices that are not likely to adversely impact students' opportunity to learn required educational content or the assessment of educators who provide instruction. This includes most research on regular and special educational instructional strategies, and the research on the effectiveness of or the comparison amount instructional techniques, curricula, or classroom management methods.**
☐	**2.**	**Research that only includes interactions involving educations tests (cognitive, diagnostic, aptitude, achievement), survey procedures, interview procedures, or observation of public behavior (including visual or auditory recording) if at least ONE of the following criteria is met:**

	☐ i.	the information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subjects;
	☐ ii.	any disclosure of the human subjects' responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation.
☐	**3.**	**Research involving benign behavioral interventions in conjunction with the collection of information from an ADULT subject through verbal or written response or audiovisual recording if the subject prospectively agrees to the intervention and information collection at least ONE of the following criteria is met:**
	☐ i.	Information obtained is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained, directly or through identifiers linked to the subjects;
	☐ ii.	Any disclosure of the human subjects' responses outside the research would not reasonably place the subjects at risk of criminal or civil liability or be damaging to the subjects' financial standing, employability, educational advancement, or reputation.
☐	**4.**	**Secondary research for which consent is not required: Secondary research uses of identifiable private information or identifiable bio-specimens, if at least ONE of the following criteria is met:**
	☐ i.	The identifiable private information or identifiable bio-specimens are publicly available;
	☐ ii.	Information, which may include information about bio-specimens, is recorded by the investigator in such a manner that the identity of the human subjects cannot readily be ascertained directly or through identifiers linked to the subjects, the investigator does not contact the subjects, and the investigator will not re-identify subjects;
	☐ iii.	The research involves only information collection and analysis involving the investigator's use of identifiable health information when that use is regulated under 45 CFR 160 and 164, subparts A and E, for the purpose of "health care operations" or "research" as defined by the regulations;
	☐ iv.	The research is conducted by, or on behalf of, a Federal department or agency using government-generated or government-collected information obtained for non-research activities, if the research generates identifiable private information that is or will be maintained on information technology that is subject to and in compliance with section 208(b) of the E-Government Act of 2002, if all of the identifiable private information collected, used, or generated as part of the activity will be maintained in systems of records subject to the Privacy Act of 1974, and, if applicable, the information used in the research was collected subject to the Paperwork Reduction Act of 1995.
☐	**6.**	**Taste and food quality evaluation and consumer acceptance studies, if:**
	☐ i.	wholesome foods without additives are consumed; or
	☐ ii.	if a food is consumed that contains a food ingredient at or below the level and for a use found to be safe, or agricultural chemical or environmental contaminant at or below the level found to be safe, by the Food and Drug Administration or approved by the Environmental Protection Agency or the Food Safety and Inspection Service of the USDA.

SECTION E: Summary

1. Please provide a ***detailed*** description of your project: **The project will be a pilot test of a virtual reality application to teach chemistry concepts. Participants will complete a pre-test, a post-test, and a post-intervention survey. The intervention itself will involve the use of virtual reality headsets and a prototype application developed by the Researcher. Participants will be given an orientation on the use of the virtual reality hardware and the application will include a brief tutorial. The project will be conducted over the course of 2-3 class periods of 55 minutes each, spanning a timeframe 3-5 days. The activity will be conducted in the Learning Commons of a public high school.**

Please provide a ***detailed*** description of your project. Include information on recruitment, consent process, incentives (if applicable), methods of data collection, data storage, etc.:

- Who will survey, interview, collect data: Data and consent will be collected by a colleague of the Researcher until all students have completed the tasks and grades have been given.

- How, specifically, are participants being recruited: Participants are high school students in their junior or sophomore year currently enrolled in a chemistry course taught by the Researcher. Students will be recruited from this sample of convenience. The project will be conducted during the course of a typical classroom period. Parental consent and participant assent will be obtained and stored by a third party, a colleague of the Researcher.

- If there is a signed agreement, who owns the data:
  ***Provide copy of agreement to IRB.**

- Will participants be offered incentives?

X NO
☐ YES, please provide details:

- If direct quotes will be included as part of the analyzed data, is this information in the consent form?

X YES* (must be included this information as part of informed consent)
☐ N/A

- Will participants be audio and/or video recorded?

☐ NO
X YES* (must include this information in the consent—include whether these recordings will be kept, or transcribed, verified and deleted so only the transcripts will be considered data.)

SECTION F: Population

1. Will your population include:

☐ Adults (18 and over)
X Minors (17 and younger): **X Ages 15-17** **☐ Ages 11-14** **☐ Ages 10 and younger**
☐ At risk population: **Explain:**

****Additional school district permissions may apply if research is being conducted in primary or secondary schools. Check with the appropriate school district(s).***

2. Will your population include Boise State University students or employees? ***If so, additional permissions may apply.***

   X NO

   ☐ YES:

   If BSU students or employees report to you, list the third party contact who will hold all data until final grades have been given or data has been coded:

SECTION G: Applicable Documents

1. Please check the boxes for the documents included with this submission:

☐ Recruitment Scripts	**X Survey(s)**	**Interview Questions**	**X Letter(s) of Support***
X Consent	**X Assent (15-17)**	**☐ Assent (11-14)**	**☐ Verbal Assent (10 and younger)**
☐ Other:			

*If you are collecting data at a location not on BSU property **OR** you are wanting to access information via your department (such as a list of email to contact all students with that major), you may also need to provide a Letter of Support.

SECTION H: Signature

My signature below and/or by submitting protocol documents from my Boise State email address indicates:

- **I agree to fully comply with the policies and procedures outlined in Boise State University's IRB Policy (5050) as well as all applicable program guides, rules and regulations.**
- **I will ensure all personnel involved in the activities outlined in this application have received training on appropriate practices and procedures.**
- **I ensure the information provided in this document is accurate and complete and that I am qualified to perform the described activities.**
- **I agree to stay within the scope of activities outlined in this application, and I understand any changes in activities must be approved by the IRB before they begin.**

PI Signature__________________________ Date_______________

Co-I Signature________________________ Date_______________

You will receive a Notification of Exemption once the IRB has reviewed this application. Research cannot begin until this letter is received.

IRB Approval

BOISE STATE UNIVERSITY
RESEARCH AND ECONOMIC DEVELOPMENT

Date: November 04, 2022

To: Brett Shelton cc: Jason Ward

From: Office of Research Compliance (ORC)

Subject: SB-IRB Notification of Exemption - 101-SB22-168

A New Reality in Teaching Chemistry: Developing a Virtual Reality Intervention for Teaching and Learning Concepts of Chemical Equilibrium

The Boise State University ORC has reviewed your protocol application and has determined that your research is exempt from further IRB review and supervision under 45 CFR 46.101(b).

Protocol Number: 101-SB22-168	Received: 9/30/2022	Review: Exempt
Expires: 11/3/2025	Approved: 11/4/2022	Category: 2

This exemption covers any research and data collected under your protocol as of the date of approval indicated above, unless terminated in writing by you, the Principal Investigator, or the Boise State University IRB. All amendments or changes (including personnel changes) to your approved protocol **must** be brought to the attention of the Office of Research Compliance for review and approval before they occur, as these modifications may change your exempt status. Complete and submit a Modification Form indicating any changes to your project.

Exempt protocols are set to expire after three years. Annual renewals are not required for exempt protocols. If the research project will continue beyond three years, a new application must be submitted for review. If the research project is completed before the expiration date, please notify our office by submitting a Final Report.

All forms are available on the ORC website at http://goo.gl/D2FYTV

Please direct any questions or concerns to ORC at 426-5401 or humansubjects@boisestate.edu.

Thank you and good luck with your research.

Office of Research Compliance

APPENDIX B

Data collection instruments

Pre-Test/Post-Test Questions

1) In an experiment, a student mixes two compounds, X and Y, which react to give off energy and form compound Z, as shown in the equilibrium equation below.

$$X + Y \leftrightharpoons Z + \text{energy}$$

After the reaction reaches equilibrium, which of the following changes would shift the equilibrium to the left?

a) adding more compound X to the equilibrium
b) adding more compound Y to the equilibrium
c) increasing the temperature of the equilibrium
d) decreasing the temperature of the equilibrium

2) According to the reaction below, increasing the temperature of the reaction would have what effect?

$$A + B + \text{HEAT} \leftrightharpoons C$$

a) Increased production of C
b) Increased production of A
c) Increased production of B
d) Equilibrium would not change

3) Consider the reaction below

$$H_2(g) + I_2(g) \rightleftarrows 2HI(g)$$

Which of the following would occur if iodine (I_2) were added to the system?

a) [H_2] decreases, [HI] increases
b) [H_2] decreases, [HI] decreases
c) [H_2] increases, [HI] decreases
d) [H_2] decreases, [HI] decreases

4) The reduction of carbon dioxide by hydrogen gas takes place at 420°C to produce water vapor and carbon monoxide. The equation for this reaction at equilibrium is shown below.

$$H\ (g) + CO_2\ (g) \leftrightharpoons H_2O(g) + CO_2(g)$$

Which of the following changes in concentration occur when more water vapor is added to the system under equilibrium conditions?

a) [H_2] decreases, [CO_2] decreases, [CO] increases
b) [H_2] decreases, [CO_2] decreases, [CO] decreases

c) $[H_2]$ increases, $[CO_2]$ increases, $[CO]$ increases
d) $[H_2]$ increases, $[CO_2]$ increases, $[CO]$ decreases

5) The equation below shows a reaction at equilibrium.

$2NO_2\ (g) \leftrightharpoons N_2O_4\ (g) + 58.8\ kJ$

What happens to the equilibrium if the temperature is increased?

a) A new product will form.
b) More NO_2 will be formed.
c) More N_2O_4 will be formed.
d) The equilibrium will remain the same.

6) In the Haber process for making ammonia, an increase in pressure favors

$N_2\ (g) + 3H_2\ (g) \leftrightharpoons 2NH_3\ (g)$

a) the forward reaction
b) the reverse reaction
c) neither reaction
d) Cannot be determined

7) For the reaction below, what would be the effect of a decrease in pressure?

$PCl_3 + Cl_2 \leftrightharpoons PCl_5 + HEAT$

a) Increased production of products
b) Increased production of reactants
c) No impact on the equilibrium
d) Cannot be determined

8) According to the reaction shown below

$$2NO(g) + O_2(g) \rightleftarrows 2NO_2(g)$$

What would the effect be of lowering the concentration of O_2 in the system?

a) [NO] decreases, [NO_2] increases
b) [NO] decreases, [NO_2] decreases
c) [NO] increases, [NO_2] decreases
d) [NO] decreases, [NO_2]HI decreases

9) The reaction below is carried out in a sealed container

$N_2\ (g) + 3H_2\ (g) \leftrightharpoons 2NH_3\ (g) + Heat$

Which of the following would increase the production of NH_3?

a) Heat the container
b) Decrease the pressure
c) Cool the container
d) None of the above

10) Given the reaction shown below, which of the following would NOT shift the equilibrium?

$BiCl_3(aq) + H_2O(l) \rightleftharpoons BiOCl(s) + 2HCl(aq) + Heat$

a) Increasing the temperature
b) Removing HCl from the system
c) Adding $BiCl_3$ to the system
d) Increasing the pressure of the system

Virtual Reality Neuroscience Questionnaire

VIRTUAL REALITY NEUROSCIENCE QUESTIONNAIRE

Please, from 1 to 7, **circle** the response that closely represents your opinion.

SCIENZA NUOVA
UNIVERSITÀ DEGLI STUDI SUOR ORSOLA BENINCASA

THE UNIVERSITY *of* EDINBURGH
School of Philosophy, Psychology and Language Sciences

User Experience

What is the level of immersion you experienced?

1	2	3	4	5	6	7
Extremely Low	**Very Low**	**Low**	**Neutral**	**High**	**Very High**	**Extremely High**

Please write below any additional comments and/or suggestions relevant to the question above:

What was your level of enjoyment of the VR experience?

1	2	3	4	5	6	7
Extremely Low	**Very Low**	**Low**	**Neutral**	**High**	**Very High**	**Extremely High**

Please write below any additional comments and/or suggestions relevant to the question above:

How was the quality of the graphics?

1	2	3	4	5	6	7
Extremely Low	**Very Low**	**Low**	**Neutral**	**High**	**Very High**	**Extremely High**

Please write below any additional comments and/or suggestions relevant to the question above:

How was the quality of the sound?

1	2	3	4	5	6	7
Extremely Low	**Very Low**	**Low**	**Neutral**	**High**	**Very High**	**Extremely High**

Please write below any additional comments and/or suggestions relevant to the question above:

How was the quality of the VR technology overall (i.e. hardware & peripherals)?

1	2	3	4	5	6	7
Extremely Low	**Very Low**	**Low**	**Neutral**	**High**	**Very High**	**Extremely High**

Please write below any additional comments and/or suggestions relevant to the question above:

Game Mechanics

How easy was to use the navigation system (e.g. teleportation) in the virtual environment?

1	2	3	4	5	6	7
Extremely Difficult	**Very Difficult**	**Difficult**	**Neutral**	**Easy**	**Very Easy**	**Extremely Easy**

Please write below any additional comments and/or suggestions relevant to the question above:

How easy was to physically move in the virtual environment?

1	2	3	4	5	6	7
Extremely Difficult	**Very Difficult**	**Difficult**	**Neutral**	**Easy**	**Very Easy**	**Extremely Easy**

Please write below any additional comments and/or suggestions relevant to the question above:

How easy was to pick up and/or place items in the virtual environment?

1	2	3	4	5	6	7
Extremely Difficult	**Very Difficult**	**Difficult**	**Neutral**	**Easy**	**Very Easy**	**Extremely Easy**

Please write below any additional comments and/or suggestions relevant to the question above:

How easy was to use items in the virtual environment?

1	2	3	4	5	6	7
Extremely Difficult	**Very Difficult**	**Difficult**	**Neutral**	**Easy**	**Very Easy**	**Extremely Easy**

Please write below any additional comments and/or suggestions relevant to the question above:

How easy was the 2-handed interaction e.g., grab the tablet with the one hand, and push the button with the other hand?

1	2	3	4	5	6	7
Extremely Difficult	**Very Difficult**	**Difficult**	**Neutral**	**Easy**	**Very Easy**	**Extremely Easy**

Please write below any additional comments and/or suggestions relevant to the question above:

In-Game Assistance

How easy was to complete the tutorial(s)?

1	2	3	4	5	6	7
Extremely Difficult	**Very Difficult**	**Difficult**	**Neutral**	**Easy**	**Very Easy**	**Extremely Easy**

Please write below any additional comments and/or suggestions relevant to the question above:

How helpful was/were the tutorial(s)?

1	2	3	4	5	6	7
Extremely Unhelpful	**Very Unhelpful**	**Unhelpful**	**Neutral**	**Helpful**	**Very Helpful**	**Extremely Helpful**

Please write below any additional comments and/or suggestions relevant to the question above:

How did you feel about the duration of the tutorial(s)?

1	2	3	4	5	6	7
Extremely More Time Needed	**Much More Time Needed**	**More Time Needed**	**Neutral**	**Enough Time Available**	**Much Time Available**	**Plenty of Time Available**

Please write below any additional comments and/or suggestions relevant to the question above:

How helpful were the in-game instructions for the task you needed to perform?

1	2	3	4	5	6	7
Extremely Unhelpful	**Very Unhelpful**	**Unhelpful**	**Neutral**	**Helpful**	**Very Helpful**	**Extremely Helpful**

Please write below any additional comments and/or suggestions relevant to the question above:

How helpful were the in-game prompts e.g. arrows showing the direction, or labels?

1	2	3	4	5	6	7
Extremely Unhelpful	**Very Unhelpful**	**Unhelpful**	**Neutral**	**Helpful**	**Very Helpful**	**Extremely Helpful**

Please write below any additional comments and/or suggestions relevant to the question above:

VR Induced Symptoms and Effects (VRISE)

Did you experience nausea?

1	2	3	4	5	6	7
Extremely Intense Feeling	**Very Intense Feeling**	**Intense Feeling**	**Moderate Feeling**	**Mild Feeling**	**Very Mild Feeling**	**Absent**

Please write below any additional comments and/or suggestions relevant to the question above:

Did you experience disorientation?

1	2	3	4	5	6	7
Extremely Intense Feeling	**Very Intense Feeling**	**Intense Feeling**	**Moderate Feeling**	**Mild Feeling**	**Very Mild Feeling**	**Absent**

Please write below any additional comments and/or suggestions relevant to the question above:

Did you experience dizziness?

1	2	3	4	5	6	7
Extremely Intense Feeling	**Very Intense Feeling**	**Intense Feeling**	**Moderate Feeling**	**Mild Feeling**	**Very Mild Feeling**	**Absent**

Please write below any additional comments and/or suggestions relevant to the question above:

Did you experience fatigue?

1	2	3	4	5	6	7
Extremely Intense Feeling	**Very Intense Feeling**	**Intense Feeling**	**Moderate Feeling**	**Mild Feeling**	**Very Mild Feeling**	**Absent**

Please write below any additional comments and/or suggestions relevant to the question above:

Did you experience instability?

1	2	3	4	5	6	7
Extremely Intense Feeling	**Very Intense Feeling**	**Intense Feeling**	**Moderate Feeling**	**Mild Feeling**	**Very Mild Feeling**	**Absent**

Please write below any additional comments and/or suggestions relevant to the question above:

Virtual Reality Neuroscience Questionnaire (VRNQ) – Scores

Section	Score	Minimum Cut-offs	Parsimonious Cut-offs
User Experience		≥ 25	≥ 30
Game Mechanics		≥ 25	≥ 30
In-Game Assistance		≥ 25	≥ 30
VRISE		≥ 25	≥ 30
Total VRNQ		≥ 100	≥ 120

The median of each sub-score and totals score should meet the suggested cut-offs to support that the evaluated VR software has an adequate quality without any significant VRISE.
The utilisation of the parsimonious cut-offs more robustly supports the suitability of the VR software.

The VR Neuroscience Questionnaire (VRNQ) was developed by Panagiotis Kourtesis in affiliation with University of Edinburgh & University Suor Orsola Benincasa of Naples.

SCIENZA NUOVA
UNIVERSITA DEGLI STUDI SUOR ORSOLA BENINCASA

THE UNIVERSITY of EDINBURGH
School of Philosophy, Psychology and Language Sciences

System Usability Scale (Items 1-10) and User Reaction Survey (Items 11-27)

Participants are asked to score the following items with one of five responses that range from Strongly Agree to Strongly disagree:

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.
11. At times I felt totally absorbed in practicing.
12. I did not find any challenge within this game.
13. Practicing this way was fun.
14. I felt engaged in my own learning while practicing.
15. Using this technology motivated me to keep practicing.
16. Practicing this way was boring.
17. I lost track of time while practicing.
18. Practicing this way was not engaging.
19. There were elements of challenge within the game.
20. I would be more likely to practice equilibrium problems this way.

21. I found practicing this way frustrating.
22. At no time was I absorbed in the game while practicing.
23. I found my way around the game easily.
24. The headgear was uncomfortable.
25. Wearing the headgear did not bother me.
26. I did not enjoy practicing this way.
27. Practicing this way will help me learn chemical equilibrium.

APPENDIX C

Full Virtual Reality Neuroscience Questionnaire Results

Full Virtual Reality Neuroscience Questionnaire Results

Survey Item	**Average Score**
What is the level of immersion that you experienced?	5.2
What was your level of enjoyment of the VR experience?	6.1
How was the quality of the graphics?	5.0
How was the quality of the sound?	4.7
How was the quality of the VR technology overall (i.e., hardware & peripherals)?	5.7
User Experience Total	**26.5**
How easy was it to use the navigation system (e.g., teleportation) in the virtual environment?	5.7
How easy was it to physically move in the virtual environment?	6.0
How easy was it to pick up and/or place objects in the virtual environment?	5.8
How easy was it to use items in the virtual environment?	5.9
Game Mechanics Total	**23.5***
How easy was it to complete the tutorials?	6.1
How helpful were the tutorials?	6.0
How did you feel about the duration of the tutorials?	5.7
How helpful were the in-game instructions for the tasks you needed to perform?	6.0
How helpful were the in-game prompts (e.g., arrows showing direction, or labels)?	6.2
In-Game Assistance Total	**30**

Did you experience nausea?	6.3
Did you experience disorientation?	6.0
Did you experience dizziness?	6
Did you experience fatigue?	6
Did you experience instability	6
VRISE Total	**30.3**
Total Overall Average	**110.3**

*** One item was not applicable to the learning environment and was removed. This reduced the minimum score threshold for this area to 20.**