An Evaluation Framework for Virtual Reality Safety Training Systems in the South African Mining Industry

by

Etienne Anton van Wyk

submitted in accordance with the requirements
for the degree of

Doctor of Philosophy

in the subject

Information Systems

at the

University of South Africa

Supervisor: Prof M.R. de Villiers

February 2015

I declare that **"An Evaluation Framework for Virtual Reality Safety Training Systems in the South African Mining Industry"** is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete references.

E A van Wyk

February 2015

Student no: 7740115

ACKNOWLEDGEMENTS

My sincerest gratitude extends to the following for their assistance during my PhD studies. Thank you to:

- My supervisor, Prof Ruth de Villiers, for the many years of guidance, advice, encouragement and support. Even through tough personal circumstances you remained committed, and I am indebted to you and extremely grateful for the many hours of your time given to assisting me, and also shaping me to be a future supervisor.

- My wife, Frenette, and my sons, Zander and Sasha. Thank you very much for your love, understanding and support during this lengthy project.

- Dr Filistea Naude at UNISA, for your assistance with sourcing relevant literature.

- Tshwane University of Technology (TUT), for granting me a period of sabbatical leave during this study.

- Students at the Centre for Creative Technologies at TUT, for the development of the prototypes.

- Impala Platinum, for granting me permission to conduct research on site in Rustenburg.

- Mrs Helene Muller at UNISA, for statistical support.

- National Research Foundation, for a Thuthuka grant to assist with the expenses of this research.

- Mrs Sheyne Ball, for your dedicated efforts in the language editing of this thesis.

Soli Deo gloria

Etienne van Wyk
February 2015

ABSTRACT

The mining industry in South Africa contributes significantly to the national economy. Despite stringent safety legislation, mining accidents cause numerous fatalities and injuries. Inadequate or insufficient training is often cited as a root cause of accidents. Conventional class-based safety training has not reduced the incidence of accidents significantly. By contrast, virtual reality training tools can provide simulated exposure to real-world working conditions without the associated risks.

This study describes the application of design-based research (DBR) in the design and development of two desktop virtual reality (VR) systems for safety training in the South African mining industry. The results of a usability context analysis were applied in the design of a VR prototype on generic hazards recognition and rectification, which was used and evaluated at South Africa's largest platinum mine site. A case study was conducted to investigate the causes and occurrences of falls of ground, which resulted in the design and development of a second VR prototype focusing on identifying and addressing underground geological conditions.

DBR was also used in the generation of an evaluation framework for evaluating VR training systems, namely the Desktop VR Evaluation Framework (*DEVREF*), which is the major deliverable of the research. *DEVREF* can make a major contribution to the domain of e-training in mines and is transferable and customisable beyond its initial application. The process flow of the research thus moved beyond merely providing a solution to a complex real-world problem and became a classic DBR study with dual outcomes, namely a *practical* real-world solution in the form of two VR training systems and a *theoretical* contribution in the form of the *DEVREF* evaluation framework. *DEVREF* evaluates the design of desktop VR training systems in the categories of instructional design, usability, VR systems design, and context-specific criteria for mining. The use of *DEVREF* is demonstrated by reporting the application of its criteria in evaluating the two VR training systems. Heuristic evaluation, end-user surveys, and interviews were used as evaluation methods.

A third contribution is *methodological*, in that this work proposes a new DBR process model and an interaction design lifecycle model suitable for VR training systems.

Keywords: design-based research, end-user surveys, evaluation framework, heuristic evaluation, instructional design, interactive e-training, meta-evaluation, mine safety training, usability, usability context analysis, virtual reality design lifecycle model, virtual reality training.

TABLE OF CONTENTS

Chapter Eight: Evaluation

Chapter Nine: Revised Evaluation Framework

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AR – Augmented Reality
CLT – Cognitive Load Theory
CoM – Chamber of Mines
CTML – Cognitive Theory of Multimedia Learning
DBQA – Database Questions and Answers
DBR – Design-based Research
DEVREF – Desktop Virtual Reality Evaluation Framework
DMR – Department of Mineral Resources
DSR – Design Science Research
FOG – Fall of Ground
GCD – Ground Control District
HCI – Human Computer Interaction
HE – Heuristic Evaluation
HMD – Head-mounted Display
ID – Instructional Design
ISGC – Interactive Simulated Geological Conditions
IVE – Interactive Virtual Environment
LSF – Look, Stop and Fix
MHSC – Mine Health and Safety Council
MHSI – Mine Health and Safety Inspectorate
MQA – Mining Qualifications Authority
MR – Mixed Reality
UCA – Usability Context Analysis
UCSD – User-centred Systems Design
UE – Usability Evaluation
VR – Virtual Reality

Chapter One
Introduction

1.1. Introduction

Mining in South Africa has been the main driving force behind the history and development of Africa's richest and most advanced economy (Coka, 2012). The South African mining industry is, however, frequently criticised for its poor safety record and high number of fatalities. Inadequate or insufficient training is often cited as a root cause of accidents (Van Wyk & De Villiers, 2009).

Virtual reality, popularly referred to as VR, is a rapidly growing technology which utilises the ever-increasing power of computing to simulate real-world and imaginary environments and situations with a high degree of realism and interaction. VR is currently being used and investigated for providing training solutions in a variety of industries. This study investigates how safety training in the South African mining industry can be improved by using VR.

This thesis reports on seven years of hands-on design, development and evaluation of innovative interventions for safety training at mines. It describes the application of four cycles of a design-based research process, which led to the implementation of two interactive desktop virtual reality training systems. Early in the design process, the need arose for appropriate evaluation methods and criteria. No single suitable evaluation framework was identified, with the result that the researcher set out to create one. Due to its important role in the proposed research, and its importance as VR training is increasingly used in the mining sector and other industrial domains, it became the primary purpose of this research. Moreover, the resultant evaluation framework can be applied not only to evaluate such systems, but its criteria can also serve as design aids.

A usability context analysis was conducted to contribute to the usability of system design. The results of the context analysis were applied in the design of a virtual reality prototype on generic hazards recognition and rectification, which was used and evaluated at South Africa's largest platinum mine site. A case study was conducted to investigate the causes and occurrences of falls of ground at the platinum mine, which resulted in the design, development and implementation of a second virtual reality prototype focusing on identification and addressing underground geological conditions. An evaluation framework for the evaluation of such systems was developed and applied to evaluate both prototypes using heuristic evaluation. The user satisfaction of both systems was also evaluated. The results of these evaluations, as well as a meta-evaluation done on the framework, then led to an improved evaluation framework.

This chapter introduces aspects of the study. Section 1.2 focuses on the background to the study, while the problem statement and research questions are presented in Sections 1.3 and 1.4 respectively. Section 1.5 describes the rationale behind the study from three different perspectives: mine safety training, educational aspects, and human computer interaction. The value of the research is discussed in Section 1.6. A brief outline of the literature studies is presented in Section 1.7. Section 1.8 explains the research design and methodology, and presents the research strategy. The scope of the study is discussed in Section 1.9, with Section 1.9.1 focusing on the domain and context, Section 1.9.2 presenting the delimiters and limitations, and Section 1.9.3 discussing the assumptions that underlie this research. Ethical considerations are explained in Section 1.10. Section 1.11 outlines the structure of the thesis and is followed in Section 1.12 by a summary of the chapter.

Figure 1.1 graphically indicates the layout of this chapter.

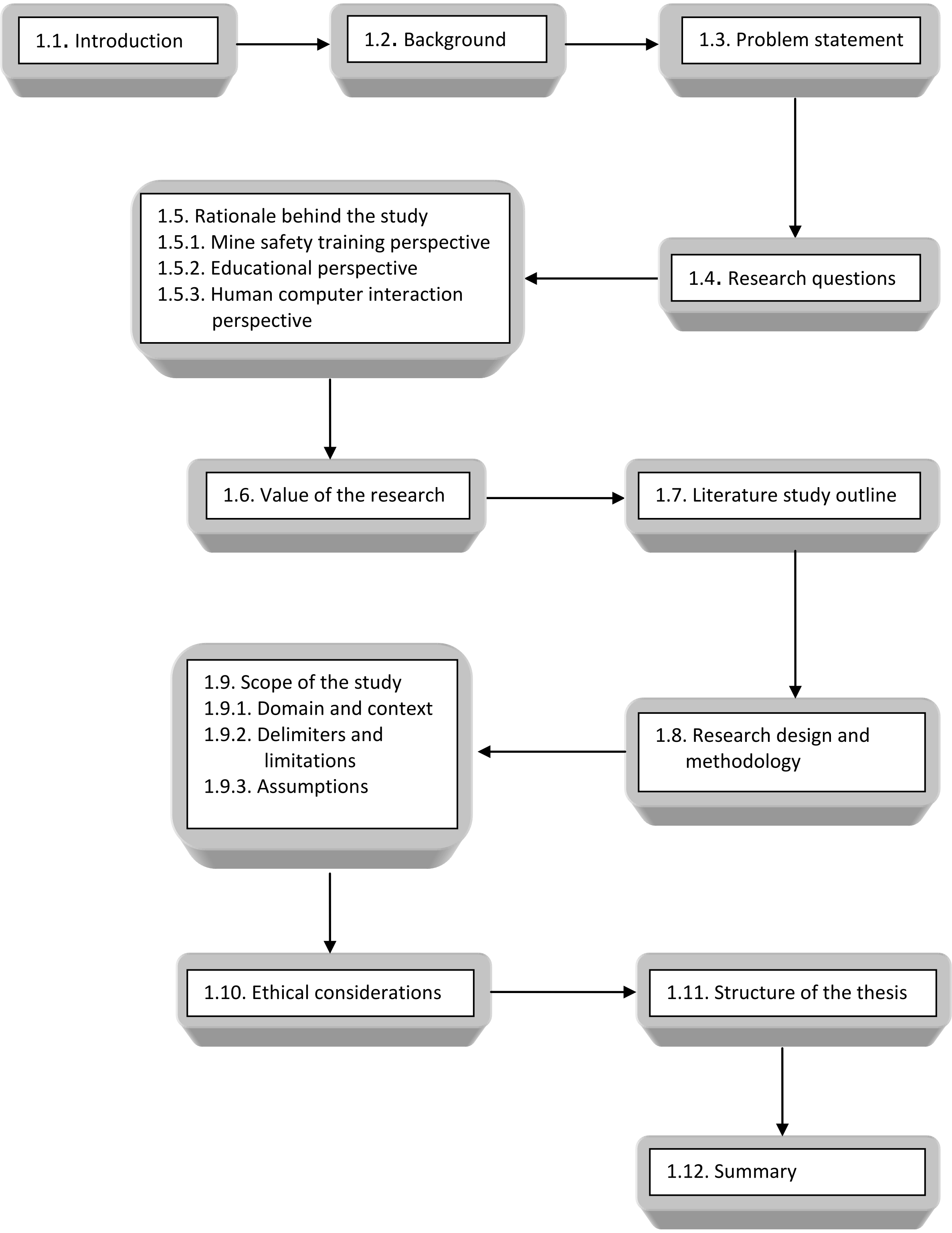


Figure 1.1: Layout of Chapter One.

1.2. Background

South Africa is well-known for a diverse wealth of minerals and a large mining industry. The mines employ hundreds of thousands of mine workers from all over Africa. The South African mining industry currently represents 20% of the country's gross domestic product (Matthee, Henneke & Johnson, 2014).

1.2.1. Incidents, accidents and fatalities

The South African mining industry, however, is also known for its high injury and fatality rate. During the previous century, over 69 000 mine workers died and over a million have been seriously injured as a result of incidents in South African mining operations (Barry, 1995:65; Krige, 1995:9). In 2003, the mining sector signed an agreement with the South African government to bring fatalities down by 20% a year in order to reach levels comparable to those of mining companies in Australia, Canada and the US. The death toll from mining accidents was 270 deaths in 2003 (CoM, 2007).

The greatest impact of mining accidents is on the victims and their family members. Many of the mine workers are the breadwinners in extended families who depend on these workers for their daily living. The impact is also felt by the mining companies and other employees, who face the constant threat of mine closures and job losses. Mining companies suffer production losses after fatalities due to routine shutdowns ordered by the government for investigations, and work stoppages by union members who stop work for a day to mark the death of colleagues.

Safety performance data for South African mines is published by the Mine Health and Safety Inspectorate. The Inspectorate also lists major accidents where four or more people were killed. There has been an average of seven such major mining accidents a year over the 20 years from 1985 to 2004 (DME, 2005). Annual fatalities decreased gradually to 199 in 2006, but in 2007, 221 fatalities were reported and in the period 2008–2009 the number of accidents at platinum mines alone increased by 18% from 1053 in 2008 to 1243 in 2009 (Citizen, 2009). More details on accident statistics are provided in Section 4.4 in Chapter Four.

In their Annual Report for 2004, the South African Chamber of Mines stated that the industry safety target is a zero fatality and injuries rate and a milestone was set to achieve constant and continuous improvement by 2013 "equivalent to current international benchmarks, at the least". All the stakeholders shared the view that safety performance had to be improved, and that “even one fatality is one too many” (CoM, 2004:107). The South African mining industry also participated in the drafting of the International Labour Organisation's Convention 176 (Safety and Health in Mines) in 1995, and the South African government ratified the convention in June 2000 (Mining Weekly, 2002).

Research into reducing subsurface mining accidents has traditionally focused on reducing fall-of-ground accidents by providing improved support units and systems and improved mining layout design (Squelch, 2000). An additional approach promoted in the Mine Health and Safety Act 29 of 1996 is, however, to improve the level and effectiveness of training given to underground workers (DMR, 2010a). In the context of underground accidents, the emphasis of this training lies in the area of hazard identification and associated remedial action.

1.2.2. Safety training

Inadequate or insufficient training is often cited as a root cause for many mining fatalities (Orr, Filigenzi & Ruff, 2002; Tichon & Burgess-Limerick, 2011). Training outside the direct working environment provides only limited real-life opportunities. As a result, such training may fail to make a significant impact in the tense working environment itself. Virtual reality-based training tools, however, can provide workers with simulated working conditions in a virtual environment, without the associated risks of the real environment. Trainees can interact with the virtual environments via a variety of hardware devices (e.g. joysticks and gloves). The impression of actually being in the virtual environment (immersion) can be created and enhanced by special optical and audio devices (e.g. head-mounted displays and 3D sound).

VR is currently being used and investigated for providing training solutions in a variety of fields such as the military, medical, power generation and aircraft industries. VR has a number of features that appear well suited to training for a mining environment and, in particular, for hazard recognition and associated remedial safety action. The primary features of relevance are: the facility to expose trainees to simulated hazardous situations without putting them in any actual danger; the facility to simulate hazardous situations more frequently than would be encountered in the real world; and the simulation of situations that have not previously occurred but which could be encountered in the industry.

Virtual reality has evolved considerably over the last two decades. "Although VR is still maturing as a technology, implications for its future as a tool for education, science, medicine and other fields, seem certain" (De Strulle, 2004:76). A major goal of this research is to explore the design and development of cost-effective virtual mine environments. These environments can be used to train underground mine workers in hazard recognition and correct safety procedures.

1.3. Problem statement

The importance of improved training to address the problem of mine safety has been noted for the past two decades. In the 1990s, studies by the US National Institute for Occupational Safety and Health (NIOSH) and the US Department of Labor indicated that mining often had the highest annual fatality rate of any private industry (Bureau of Labor, 1999). In July 1999, the Mine Safety and Health Administration in the US launched a special training initiative to prevent mining accidents, stating that training played a vital role in preventing deaths, injuries and illness on the job. "Only with effective training can miners recognise possible hazards and know the safe procedures to follow" (MSHA, 1998:3).

In South Africa, a research study conducted by the National Productivity Institute on the identification or causes of roof or sidewall accidents, identified outdated training methods and materials as a major contributing factor (Hamilton-Atwell, Du Toit, Kirstein, Louw, Mtombeni & Moses, 1997).

Legal requirements

The Mine Health and Safety Act (MHSA, 1996) has identified the creation of a culture of health and safety as one of its objectives**.** This objective is supported by a number of statutory provisions which require employee participation, instruction and training of employees, risk management, disclosure of information to employees and the employee's right to leave any work place which poses a serious danger to health and safety. The Act requires the employer to consider, as far as reasonably practicable, an employee's training and capabilities in respect of health and safety before assigning a task to that employee. In addition, the employer is required to provide employees, as far as reasonably practicable, with any information, instruction, training or supervision that is necessary to enable them to perform their work safely and without risk to health (Le Roux, 2005).

Whilst employers are primarily responsible for providing safe and healthy workplaces, the Department of Minerals and Energy is the lead agent in promoting, monitoring and enforcing legislation and initiating prosecution in terms of the Mine Health and Safety Act of 1996 (MHSA, 1996). Guidelines regarding the enforcement of the Act were released in January 2005 by the South African Government. This document contains instructions enforcing compliance with any provisions of the Act and makes provision for fines and prosecution of offenders (DME, 2005).

Mine safety

Despite the stricter enforcement of the Act, the safety record of South African mines did not show major improvement. This prompted the then Minerals and Energy Minister, Buyelwa Sonjica, to comment on this situation during June 2007: "The South African government plans to deal severely with mining companies operating in the country if their safety records do not improve" (Mining Weekly, 2007:1). She stated that CEOs of mining firms should commit more to the safety of their work force, and urged top executives to show more visible leadership. This came in the light of the fact that the industry's safety performance had not improved despite the 20% target set in 2003.

After 3 200 miners were trapped 2.2 km underground at Harmony Gold's Elandsrand Mine near Carletonville in October 2007, a national safety audit was requested by the then president of South Africa, Thabo Mbeki. Furthermore, the president requested an audit of all the mines to determine whether they met health and safety standards as prescribed in South African law (News24, 2007).

In 2008, with the 2007 fatalities recorded as 221, government started to intervene in operations of individual mines. When serious accidents occur, the mine or a particular mine shaft is closed for investigation. Mines cannot continue production until the mine is certified compliant with all safety standards. Such closures have led to major losses in production. In January 2009, the chief executive of one of South Africa's largest gold producers, Gold Fields, indicated in a press release that the group had lost R2.29 billion in revenue as a result of safety stoppages in the previous year from January to June – when a record 47 people died at Gold Fields' mines – and R290 million in the six months to December, when eight people died (Business Report, 2009). In the half-year to December, AngloGold Ashanti had nine mine deaths, while Harmony Gold had eleven.

In 2014, after the mining industry failed to reach the 2013 milestones as agreed with government, Mineral Resources Minister Susan Shabangu stated that although the mining industry had been the backbone of the South African economy and a major provider of employment, “the benefits of these contributions to development have always been overshadowed by the continued loss of life, occupational diseases and injuries” (Odendaal, 2014:1). She also expressed the need for further research on ways to improve safety and safety training, indicating that although technology had advanced, accidents were still occurring.

Training

Whilst new training rules and regulations have been enacted, many training tools and techniques are less effective than they could be in providing safety training. Prior to this research, meetings of Safety, Health and Environment managers at South African mines indicated that the mining community required improved training tools (Baker, 2006; Moldenhauer, 2006; Wenhold, 2006). During interviews conducted at two large South African mines, the mine managers specifically requested help in the development of new safety training methods. They mentioned the importance of effective training and the need for improved and updated training (Lubbe, 2006; Stander, 2006).

Current training methods used in mines rely mainly on repetitive classroom-style learning, with some instruction being given in a physical mock-up of an underground workplace followed by on-the-job training. However, under classroom conditions, workers do not make safety decisions under the same situation of stress they would experience while underground. As a result, depending on the stress levels of the real working environment, the decisions that are taken in the authentic underground environment may differ significantly from those taken under more relaxed circumstances. To enhance the effectiveness of training, an alternative training design is required that simulates the real threats as closely as possible. Squelch (2001) indicated that research into the reasons for fall-of-ground accidents in the South African mining industry highlighted shortcomings in the conventional training approach. These shortcomings indicated an opportunity for an innovative approach to be taken to improve safety and hazard awareness training.

Virtual reality and mining

Virtual reality systems range from tactile systems that physically represent the real world through to purely computer-generated visualisations. These computer-generated, three-dimensional, artificial worlds are commonly referred to as *virtual environments*

(VE), and in many cases users are able to interact with the data and images that are presented by these computer-based visual systems.

In a mining context, a primary aim of developing virtual environments is to allow mine personnel to practise and experience mine situations, activities and processes that can be encountered in the day-to-day operations at a mining site. Safe and efficient planning and production are fundamental to profitable mine operations and VR provides an intuitive means of exploring the diverse and disparate information associated with mining processes.

VR has already been shown to be an effective training tool in many industries. "The general belief is that the information and skills acquired using VR training transfer to the real world in a more meaningful and realistic way than the information and skills acquired using more conventional, didactic training methods" (Filigenzi, Orr & Ruff, 2000). VR offers the potential to expose personnel to simulated hazardous situations in a safe, highly visual and interactive way. Customised simulations of mine layouts and comprehensive virtual environments can be set up allowing users to move around the virtual mine and to take decisions. The consequences of both correct and incorrect decisions can be immediately fed back to trainees, giving them the opportunity to learn directly from their mistakes. In addition, VR allows the trainees to experience conditions that would be difficult or impossible to re-create in the real world. VR simulations can provide a wide range of possible training scenarios without incurring the high costs and risks of personnel and equipment.

This study thus proposes the design, development and implementation of interactive virtual reality training systems as an innovative approach to improve safety training. However, since this study falls within the domain of information systems, it is of particular importance to determine the effectiveness of the design of such systems, hence an approach is required that evaluates the appropriateness and effectiveness of the VR systems design within the context of mine safety training. Therefore, this study proposes the development of an evaluation framework for this vital purpose.

Cross-disciplinarity

The research problem of this thesis is, by nature, cross-disciplinary, involving activities, simulation and modelling that cut across a broad range of computing and industry-related fields. These include the design and evaluation of virtual reality software and educational computing, as well as the attainment of usability. The research uses experience from other contexts such as the military, medical training, power generation and aircraft industries, where VR training has been successfully applied. Examples of the use of VR in training are supplied in Section 2.4.

1.4. Research questions

The main purpose of this study is to present an evaluation framework for virtual reality training systems for the South African mining industry, and to demonstrate its application in evaluating two prototype VR training systems. The secondary purpose of the research is to produce novel e-training interventions on the topic of safety training for mine workers operating in the underground mining environment. The training is to be delivered by desktop VR technology.

The main research question addressed by this study, is:
What is an appropriate and effective framework for evaluating virtual reality training systems in the mining industry?

In order to address the main research question, six research subquestions are defined, as indicated in Table 1.1.

Table 1.1. Research subquestions of the study.

#	Research subquestion
RQ1	What is the suitability and potential of virtual reality technology for training applications in the domain of mine safety training?
RQ2	Which research paradigm is appropriate for the intended research?
RQ3	What are the contextual requirements for virtual reality training systems for the mining industry?
RQ4	What is an appropriate design lifecycle model for interactive desktop virtual reality training systems?
RQ5	What structure, categories and criteria should be incorporated in an evaluation framework for virtual reality training systems in the mining industry?
RQ6	How appropriate and effective is the proposed framework?

Table 5.1 in Chapter Five revisits the research subquestions and indicates in which section(s) in the study each subquestion is addressed. Chapter Ten, the concluding chapter, revisits the research subquestions to summarise their answers.

1.5. Rationale behind the study

The rationale for this study is described from three perspectives.

1.5.1. Rationale from a mine safety training perspective

During 1996, a research study was undertaken by the South African Safety in Mines Research Advisory Committee (SIMRAC) on human computer interaction in rock engineering. This study proposed the use of virtual reality simulation as an additional training method and concluded by stating that "the indications are that VR training simulators are an appropriate way of training underground workers". The study recommended that "for the implementation of VR as a mining industry training tool, fully-featured VR simulators will need to be constructed" (Squelch, 1996).

At the commencement of this study, very few studies on VR in mining could be found in literature, with only one in South Africa. Squelch (1998) developed a prototype of a virtual stope panel section of a gold mine and compared this VR method to video-based training. Structured interviews were used to determine the opinions of some mine workers at the Elandsrand Gold Mine and the study concluded that the miners strongly favoured the future implementation of VR training simulators.

According to Squelch (2005), in work that followed his earlier studies, a CSIR company, Miningtek, was planning to do further research on utilising VR for training and developing VR-based simulators, but this did not materialise due to:

- scepticism as to VR's suitability as a viable and better medium of training;
- lack of funding;
- the diverse skills required for the development team, i.e. training/instructional design specialists, programmers, graphics artists/modellers and mine safety experts;
- perceived lack of realism;
- the need for a more intuitive/natural user interface; and
- Squelch having emigrated from South Africa to Australia.

When this research commenced in 2006, as far as could be determined, the only VR-based training tools used in the South African mining industry were a coal cutter operator training system, which was used by Sasol, and a truck driver simulator used at Kumba Resources' iron mines. VR was not being used for safety training.

"Virtual reality provides the best tools for accident reconstruction, training and hazard identification by immersing the trainee in an environment as close to real world as possible" (Orr, Filigenzi & Ruff, 2002). The use of high quality three-dimensional graphics, sound and dynamic simulation can be combined to form a uniquely engaging experience. Through safety, visualisation and education, VR can provide many improvements for the minerals industry. As with other e-learning and e-training products, VR systems have the advantage that they can be used in a flexible way not restricted to prescribed sessions in a classroom, but evaluations should be conducted to determine their strengths and inadequacies.

Advances in VR technology mean that it is now both feasible and cost effective to consider mass training of workforces using simulated computer representations of the workplace. The indications, therefore, seem to be in favour of VR having meaningful application to mine safety training, and make it worthwhile investigating its application for hazard awareness training.

1.5.2. Rationale from an educational perspective

Some narrow definitions for e-learning define e-learning exclusively as using the Internet for instruction and learning, but other definitions are broader. Sangrà, Vlachopoulos and Cabrera (2012) define e-learning broadly as an approach to teaching and learning that is based on the use of electronic media and devices as tools for improving access to training, communication and interaction, and that facilitates the adoption of new ways of understanding and developing learning. It follows then that the VR training solutions proposed by this study can be viewed as e-learning artefacts.

E-learning applications reflect different views on cognition and learning, including behaviourism, cognitivism and constructivism. These learning perspectives provide structured foundations for planning and conducting instructional design activities and are discussed in more detail in Section 3.4 in Chapter Three.

E-learning applications should support learners in the process of learning. According to De Villiers (2005), this process involves information transfer, management of educational interaction, the support of human cognition, implementation of behavioural change, and leveraging technology as a medium or messenger and not as a message in its own right. "Foundations for e-learning must be based on sound principles of learning theory and instructional design, in order to facilitate effective learning" (De Villiers, 2005:351). Reigeluth (2013) describes instructional design as involving methods of instruction and contextualisation to the situations in which those methods should be used. The VR prototypes proposed by this study should therefore be designed while considering instructional design principles, including cognitive load theory, cognitive theory of multimedia learning and instructional design principles for multimedia. These aspects are addressed in Section 3.5, and are included as criteria in the evaluation framework proposed by the study.

1.5.3. Rationale from a human computer interaction perspective

Human computer interaction (HCI) is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and the study of major phenomena surrounding them (ACM, 1996). It follows then, with this study focusing on the design, evaluation and implementation of VR training systems, that aspects of HCI are of particular importance.

The foundations of HCI focus on the psychological and physiological attributes of the human user, the capabilities and limitations of computing devices, and the dialogue between the two. HCI design practice addresses usability from the human perspective. The International Organisation for Standardisation defines usability as the "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO 9241-210, 2010). Usability plays an important role in the success of e-learning applications. If an e-learning system is not usable, the learner will be spending too much time on understanding the software functionality rather than understanding the learning content (Costabile, De Marsico, Lanzilotti, Plantamura & Roselli, 2005).

General usability and human factor issues of VR have been examined by several authors, including Galimberti and Belloni (2003), Sutcliffe and Kaur (2000), Wilson (1997) and Bowman, Gabbard and Hix (2002). Further investigation is required regarding aspects of participation in virtual environments and their consequences for usability that is specific to workplace training.

Evaluation is required to verify that users can use the product and that they enjoy using it, particularly if the design concept is new. Evaluation is concerned with gathering data about the usability of a design or product by a specified group of users for a particular activity within a specified environment or work context (Rogers, Sharp & Preece, 2011). Since this study proposes novel interactive training simulations using virtual reality, evaluations of these systems are imperative to assess their value regarding effectiveness, efficiency and user satisfaction. The evaluations should be performed based on criteria relevant to the context of virtual environments.

A number of authors stress the importance of context in designing and evaluating educational software (Jones, Scanlon, Tosunoglu, Morris, Ross, Butcher & Greenberg, 1999; Mayes & Fowler, 1999; Squires & Preece, 1999). These authors also show that there is scope for identifying synergies between usability and educational computing. In fact, there should be a synergy between the learning process and the interaction with the application (Ardito, Costabile, De Marsico, Lanzilotti, Levialdi, Roselli & Rossano, 2006). This supports the notion of integrating usability with learning.

Section 1.3 highlighted the need for improved training and an alternative means of delivering it, hence the proposed development of two interactive VR prototypes. The need now arises for a custom-built means of evaluating such systems with their unique requirements and underground context.

According to Rogers *et al.* (2011), the three main evaluation approaches for interactive systems are usability testing, field studies and analytical evaluation. Various factors impact on determining suitable theoretical foundations for e-learning applications. No single paradigm is appropriate for all situations, since domain, context and content will have to be considered (De Villiers, 2005). Technological issues and educational theories should be considered to find an appropriate solution.

Tsiatsos, Andreas and Pomportsis (2010:67) point out that "there is a need for a detailed theoretical framework for VR-based learning environments that could guide future development efforts". They then propose a framework for a specific category of collaborative virtual environments, which entails group work in a single immersive environment. This evaluation approach is not appropriate to the present study, which relates to individualised desktop VR training. Hanna, Nader and Richards (2014) propose an evaluation framework for virtual reality, but this framework similarly focuses on collaborative virtual environments, which fall outside the scope of this study.

Some earlier frameworks include the work by Bowman, Koller and Hodges (1998), who presented a framework for the analysis and evaluation of travel techniques in immersive virtual environments. However, this framework is limited to viewpoint motion control techniques used in such environments. Bowman, Gabbard and Hix (2002) presented an overview of usability evaluation of virtual environments and only discuss issues that differentiate usability evaluation of virtual environments from evaluation of traditional user interfaces. Gang, Jun and Yingzhen (2006) proposed an evaluation framework for the evaluation of virtual geographic environments, but this framework only evaluates three aspects, namely the reality portrayed, immersion and usability of such environments.

Current evaluation frameworks are limited, because they are either confined to evaluation of a specific type of virtual environment or they focus on a restricted aspect of virtual environments. This study addresses the gap for a framework for evaluation of desktop VR training systems for the mining industry, by investigating the design and development of such systems meticulously and comprehensively from the following perspectives: instructional design, usability, and VR systems design, situated in the context of underground mining. These different perspectives are integrated into a single framework, providing a multi-faceted evaluation approach.

1.6. Value of the research

According to the South African Mine Health and Safety Act (MHSA, 1996), the employer is required to ensure that every employee becomes familiar with work-related hazards and risks and the measures that must be taken to eliminate, control and reduce those hazards and risks. It must also be borne in mind that the employer is ultimately responsible for ensuring that all employees, including contractual employees, are properly trained.

Despite this legislation, mine fatalities and injuries have reduced only marginally during the last few years. Accident statistics show that human behaviour is the primary cause of mining accidents (Le Roux, 2005). This emphasises the importance of safety training. “Deficiencies in risk assessment, hazard monitoring, medical surveillance and training, along with limited access to technical expertise and management systems, constitute a significant and systematic root cause of risk” (Biffi, 2000).

As a result of the commitment to safety training by the South African Chamber of Mines (CoM, 2004) and at the request of the mining industry for more effective, engaging training tools (Baker, 2006; Lubbe, 2006; Stander, 2006; Van der Sandt, 2009), this research will propose new training methods for safety training which will result in the development of accessible and affordable VR training software. The software will be used to help reduce the fatalities and serious injuries associated with mining accidents.

VR technology has developed rapidly and costs have fallen to levels where it can now be considered for mainstream training applications. The availability of 3D modelling tools and simulation programming engines that work effectively with a mid-range desktop PC and a standard 3D graphics card make VR even more attractive to mine training centres (Van Wyk & De Villiers, 2009).

Virtual reality offers notable possibilities in training, simulation and education. Simulations can be developed for a particular situation and can often be modified for other similar situations. Although the minerals industry has been slow to invest in, and use, this advanced technology, the number of VR applications in the industry internationally is increasing. VR has a great potential to increase productivity and better utilise time. Most importantly, it can improve safety awareness and therefore reduce incidents (Stothard & Swadling, 2010).

The evaluation framework presented by this study can be used as a set of design principles to inform design of VR training systems for the mining industry, or as an evaluation tool comprising criteria/heuristics to assess effectiveness of the design of such systems, specifically relating to usability, instructional design, VR systems design and mining industry context-specific aspects.

An unanticipated contribution of this study is the use of the developed prototypes for training mining engineering students at the Universities of Pretoria and Johannesburg. Many students, even in their second year of study towards a mining engineering qualification, have never been underground in a mine and the VR prototypes provide a realistic view into the underground environment.

A further result of this study is the commitment from management at various mines to invest in computer training facilities. In cases where these prototypes were deployed, the mines first had to purchase computers before they could implement these training systems. The resultant computer training facilities can also be used for other e-learning programs.

Potential beneficiaries of this study are the more than 200 000 mine workers who work underground in South African mines. To date, interactive training systems resulting from prototype systems described in this study, have been implemented on fifteen training centres at various mines and smelting plants, and these mines are committed to further development of VR training systems. At the mine where the empirical work of this study was done, more than 17 000 employees are undergoing this training annually.

The international mining community can also benefit from this study. The researcher is a member of the International Group for VR in Mining, headed by Dr Phil Stothard from Australia, and was involved in co-authoring a collaborative research paper with seven other members of the group. It was presented at the Future Mining conference in Australia (Stothard, Squelch, Van Wyk, Schofield, Fowle, Caris, Kizil & Schmid, 2008).

As a result of this study, the researcher has authored or co-authored eleven conference papers (of which one has been cited more than 50 times by other researchers on virtual reality), and one journal article published in 2013. The conference paper on *Incident Reconstruction Simulations – its potential impact on the prevention of future mine incidents*, co-authored by the researcher, received the gold medal of the South African Institute for Mining and Metallurgy for best paper of 2011 in the mining industry. Selected papers are included in Appendix D.

1.7. Literature study outline

Due to the interdisciplinary nature of the research, this thesis has three literature reviews:

- Chapter Two describes the application domain of the research, namely *virtual reality* (VR). It reports on a study of literature on VR and its applications, with specific reference to the mining industry. Various definitions of VR are presented and the categories and features of VR systems are discussed. A classification of available systems is presented to alleviate potential confusion when developing VR applications. The application of virtual reality technology is discussed and several examples are presented where VR technology has been developed for research and industrial applications. The use of virtual reality for training purposes and, in particular, training in the mining industry is also covered.

- Chapter Three discusses human computer interaction (HCI) aspects relevant to this research, namely systems design, usability and instructional design. As this study focuses on VR training applications, the chapter also considers relevant learning theories. Furthermore, the chapter takes an in-depth look at instructional design, with subsections covering the psychological theory underpinning design, design of multimedia learning, and methodologies that facilitate learning. Lastly, the chapter reports on various usability evaluation methods and concludes with heuristic evaluation, one of the primary research methods of this study.

- Chapter Four focuses on current safety practices in the South African mining industry. Despite stringent safety legislation, accidents in the mining industry are still causing high numbers of fatalities and injuries. This chapter moves beyond the literature and lays the foundations for the empirical work of this research, by relating the theory and legislation to application. An introduction to the South African mining environment is followed by an overview of the safety legislation applicable to the mining industry. An overview is provided of the major stakeholders involved in the industry, and mine safety statistics are presented and discussed. Each major section in the chapter ends with a subsection called *Application to Training*. These subsections link the topics to the main focus of the study, namely the improvement of mine safety training using virtual reality.

1.8. Research methodology

This study describes the application of design-based research in the design and development of desktop virtual reality training systems, and the generation of an evaluation framework for such systems. This research is predominantly a quantitative study, but also has a qualitative component via interviews and some open-ended questionnaires. Other research methods applied are prototyping, surveys, heuristic evaluation, a case study and informal participant observation. These approaches and methods are considered in the sections that follow.

1.8.1. Research paradigm

The underlying research paradigm of this study is *design research,* which is currently a maturing research methodology within a number of disciplines. Design research originated from the work of Simon (1981), who distinguished between the *natural sciences*, such as anatomy, astronomy, chemistry and physics, and the sciences of the artificial, or *design sciences,* such as engineering, product design, information technology and instruction*.* In the natural sciences, *descriptive* laws represent natural phenomena, while theories and formulae explain how they occur. The design sciences, by contrast, relate to man-made phenomena, where theories and models outline goals to be achieved and procedures to accomplish them, which are set out by *prescriptive* laws. Design science is characterised by the construction and evaluation of innovative artefacts and interventions in authentic settings.

Applied design science led to design research, which is called *design science research* (DSR) in the discipline of information systems and *design-based research* (DBR) in the fields of educational technology and e-learning.

Design research is increasingly used for studies within the context of educational technology, especially for studies on the development of e-learning and e-training (De Villiers, 2012). The term used to describe design research in this context is design-based research. Design research, in particular design-based research, was selected as the underlying research paradigm of this study because of its cyclic nature of design, evaluation and redesign, and its mandatory production of both theory and actual solutions in real-life contexts, in this case, the context of instructional system design. An alternative paradigm could have been action research (De Villiers, 2005b; McNiff, 2013; Noffke & Somekh, 2009), since it is also iterative and can apply to inventions, interventions and products. However, DBR was deemed the most appropriate choice for this research due to its focus on:

(i) solving complex problems,
(ii) producing authentic artefacts, and
(iii) generating dual outcomes.

An alternative, more recent term for DBR used by some researchers is 'Educational design research' (Plomp, 2007; Teräs & Herrington, 2014; Van den Akker, Gravemeijer, McKenney & Nieveen, 2006), but the present study uses the term 'design-based research' throughout.

Design-based research is systematic and flexible, and aimed at improving educational practices (MacDonald, 2008). DBR methods attempt to bridge theory and practice in education as they uncover relationships between educational theory, designed artefacts, and practice (Design-based Research Collective, 2003). DBR is a pragmatic approach that can improve educational research to yield discernable benefits and impact on practitioners, while also being socially responsible (Reeves, Herrington & Oliver, 2005).

1.8.2. Research methods

The research methods used in this study are: literature studies, prototyping, survey research, case study research, informal participant observation, and heuristic evaluation. The main strategies are the methods employed in evaluating the prototypes, namely heuristic evaluation and survey research, while the other methods are applied in supporting roles. Due to the study's focus on evaluation, more than one method is used to strengthen the process, namely an end-user survey and an expert evaluation method.

1.8.2.1. Literature study

A literature survey usually starts with a review of the literature dealing with the chosen topic. This sets the scene for a clear formulation of the research problem (Welman & Kruger, 2001). The literature review provides secondary data and reveals inconsistencies and gaps that may justify further research. It also enables researchers to indicate exactly where the proposed research fits in. It brings the reader up to date on previous research and related work in the areas relevant to the study, and can also point out agreements and disagreements among previous researchers (Babbie, 2010). During the course of the research project a comprehensive series of literature reviews was conducted to provide both background to support, and a foundation on which to build, the resultant training applications and evaluation framework.

1.8.2.2. Prototyping

The term *prototype* refers to a "simplified program or system that serves as a guide or example for the complete program or system" (Olivier, 2004). Though programming *per se* is not research, prototyping can be applied to demonstrate that a new model or method can indeed be implemented. Prototypes serve as vehicles for experimentation and the construction of the prototype can also provide new insights. For this research, prototyping is important to prove implementation of the proposed training interventions and for users to be able to evaluate the systems, so that they can be improved, corrected and refined.

1.8.2.3. Survey research

According to Babbie (2010:252), survey research is "probably the best method available to the researcher who is interested in collecting original data for describing a population too large to observe directly", and "surveys are also excellent vehicles for measuring attitudes and orientations in a large population". These characteristics make survey research ideal for this research among end-users in this study, where questionnaires and semi-structured interviews were used to collect data at several stages of the study, among varying groups of participants.

1.8.2.4. Heuristic evaluation

Heuristic evaluation is an inspection technique whereby a small number of experts apply a set of usability principles called *heuristics*, to evaluate whether a user interface conforms to these principles (Madan & Dubey, 2012; Zaibon and Shiratuddin, 2010). According to Rogers *et al*. (2011), the way in which experts are intended to use these heuristics is by judging them against aspects of the interface. Heuristics are usually derived from academic and professional research studies, existing criteria lists, or field observations and prior experience in the given domain (Karoulis & Pombortsis, 2003). In this study, a questionnaire was completed by the expert evaluators after they had worked through the prototypes they were evaluating.

1.8.2.5. Case study research

A case study investigates a contemporary phenomenon within a real-life context and can provide qualitative and/or quantitative data (Olivier, 2004). A case study is often used in such a situation to explain causal links in real-life situations when it is difficult, complex or impossible to use other research methods such as experiments (Gillham, 2000). In such cases, the data obtained would be more comprehensive than that obtained from a survey among a sample of the population. Due to falls of ground being the greatest contributing factor to mining injuries, a case study was used in this study to analyse the circumstances relating to fall-of-ground incidents. The findings of the case study informed the design of the second prototype.

1.8.2.6. Informal participant observation

Participant observation requires the researcher to take part in, and report on, the experiences of the members of a group, community or organisation involved in a process or event (Welman & Kruger, 2001). The participant observer becomes a member of the group or event being studied, in order to personally experience what the group members experience, understand their environment, and comprehend the meaning and significance of their behaviour. The researcher therefore performs a dual role of experiencing the activities of the group and also observing and recording such experiences. In this study, informal participant observation was done by the researcher during the usability context analysis described in Chapter Six, as well as during the evaluation of trainees using the first prototype, where it became evident that trainees' inexperience in using computers caused them to struggle in interacting with the system.

1.8.3. Research Strategy

The research strategy involves the following actions:

- Define the problem relating to mine safety.
- Do literature studies on the three theoretical focus areas of the study: VR, HCI and Mine Safety Practice.
- Define and motivate the research methodology to be used.
- Perform a usability context analysis to inform the design of the first prototype.
- Design, develop and perform a basic evaluation on the first prototype.
- Improve the first prototype based on the evaluation results.
- Conduct a case study to analyse the circumstances relating to fall-of-ground incidents, which informs the design of the second prototype.
- Design and develop the second prototype.
- Develop the theoretical evaluation framework.
- Develop the user satisfaction questionnaire.
- Evaluate both the improved version of the first prototype and the second prototype, by applying two evaluation methods: heuristic evaluation, using the evaluation framework; and surveys, using the user satisfaction questionnaire.
- Improve the prototypes based on the evaluation feedback.
- Develop a meta-evaluation instrument.
- Perform a meta-evaluation on the evaluation framework.
- Refine the evaluation framework based on the feedback of the evaluation of the prototypes and the meta-evaluation.
- Write up recommendations and future work.

The research strategy is illustrated in Figure 1.2. The actions indicated in gold colour contributed to the practical outcome of this study, namely the two VR training prototypes, whereas the actions in dark blue contributed to the theoretical outcome of the study, namely the evaluation framework.

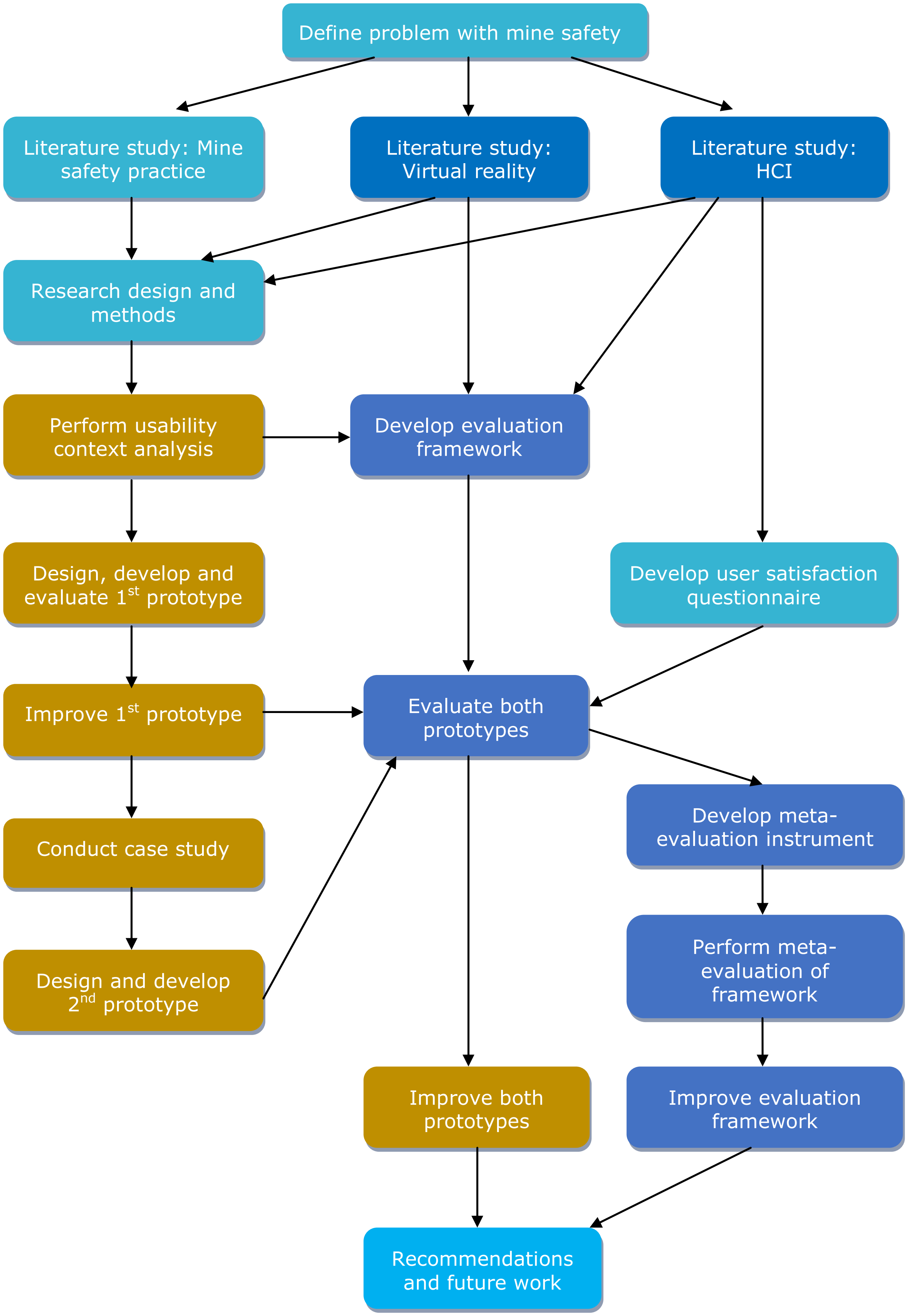


Figure 1.2: Research Strategy.

1.9. Scope of the study

1.9.1. Domain and context of the study

As indicated in the title of this thesis, the study applies to the domain of safety training in the South African mining industry. The training prototypes described in this study cover mine safety learning material, and the evaluation framework contains some criteria relevant to the mining industry.

The training content relates to generic and geological hazards in the underground mining environment. The hazards were selected from conventional, hard rock mining environments and the learning content of the prototypes do not cover mechanised mining.

The introduction of VR technology brings about excitement and high expectations of its capabilities among trainees and instructors. However, as with any other instructional media, VR should not be seen as a panacea that will work for all kinds of training content. Some learning tasks may be ideally suited for virtual representation, while others may not be effectively performed in such environments. Determining appropriateness of content to be learned with this technology falls outside the scope of this research.

The evaluation framework presented by this study is designed for heuristic evaluation, but the criteria in the framework can also be applied as design principles during the design of VR training systems. Moreover, the framework developed as a result of this research is an evaluation framework and not a conceptual framework.

It cannot be determined in this study whether the actual behaviour of the mine workers exposed to the VR training changed in practice within the underground environment. Rather, the purpose of this research, which is situated in the domain of information systems, is to investigate the developed e-training systems from the perspectives of instructional design, usability, VR systems design, and context-specific aspects relating to mine safety training.

1.9.2. Delimiters and limitations

All the empirical work for this study was done at South Africa's largest platinum mine near Rustenburg in the North West Province. This means that the learning content relates mostly to the underground environment in a platinum mine. However, some of the content is generic to all underground mining environments, and the visuals, user interaction and content in the training prototypes can be adapted and customised for training systems in other types of mining, while still using the same VR technology. The specific learning content is therefore less important to the outcomes of this study then determining effectiveness of the design of VR training systems and the application of VR technology to develop such systems.

Section 2.2.1 in Chapter Two discusses three categories of VR systems, namely immersive, semi-immersive, and desktop VR systems. Immersion is the extent to which the senses of a user perceive the illusion of reality delivered by the VR system's display. The evaluation framework presented in this research applies specifically to desktop VR. The reasons for applying desktop technology are as follows:

- Desktop VR systems run smoothly on standard desktop PCs, using input devices such as the keyboard and mouse. At most, an additional graphics card with on board memory may be required for effective delivery.
- The lower capital cost of hardware, software and peripherals make desktop VR systems an attractive and realistic alternative, available at a reasonable price.
- The idea of computer-based training was a novelty. Before the first prototype was implemented at the mine in Rustenburg, the mine was not using computer-based training at all. Ten computers had to be purchased specifically for the evaluation and were installed in a makeshift venue. After the positive results of the evaluation, mine management decided to purchase 80 computers and dedicated a venue to the VR training.
- The general low literacy levels of employees in the mining sector compounds the problem of safety training. As indicated in the analysis of the biographical information of the participants in Section 8.4.1, two-thirds of the participants had never used a computer before. VR training had to be gradually introduced to the mining industry.
- After successful deployment of the first prototype, mine management decided that all underground workers should undergo VR training. More than 17 000 miners work underground at the mine, which meant that the delivery platform of the training systems had to be suitable for large numbers of trainees. The availability of VR on relatively low-cost desktop PCs has facilitated broad use of VR technology.

The application of immersive systems is usually highly individualised since each trainee requires separate equipment to interact with the system. The high cost of such equipment makes it unattainable for training high volumes of trainees simultaneously. This study aims at improving safety training of the underground mining workforce, and with more than 200 000 miners working underground every day in South Africa, the focus is on proposing solutions that can cater for large numbers. A non-immersive training solution is therefore more viable and attractive at this stage.

The use of fully immersive systems, using cybergloves and head-mounted displays, will not form any part of this research. Therefore, the evaluation framework will not include criteria related to immersive systems, but it can be expanded to include such aspects in future, if and when more immersive systems are gradually introduced to the mining industry.

The VR prototypes in this study relate to simulated conditions in the underground environment and not the simulation of operation of equipment.

The following terms are used interchangeably in this study:

- 'participant' and 'respondent',
- 'criteria' and 'heuristics',
- 'e-learning' and 'e-training'.

1.9.3. Assumptions

Nearly all underground workers in South African mines are of the male gender, but some mines have recently started also employing female underground workers. Please note that in all places where the thesis refers to the male gender, both genders are implied.

The following assumptions were made:

- It is assumed that the questionnaires were completed by the intended persons and that such participants provided authentic and honest opinions.
- It is assumed that the research instruments yielded accurate data and the analysis of data is correct, leading to feasible recommendations.
- It is assumed that the data collected is a realistic reflection of the population being surveyed.

1.10. Ethical considerations

The miners involved in the initial evaluations were selected by the training facilitators from current trainees in their training facilities. These participants were informed in full with regard to the objectives of the evaluations. A possible ethical consideration relates to the initial withholding of a potentially beneficial training exposure from other trainees not selected for the VR evaluations. Since the mine intended implementing the system if positive results were obtained, all trainees will be exposed to the system eventually.

The researcher applied for ethical clearance from the Ethical Clearance Subcommittee of Unisa's College of Science, Engineering and Technology. This application included details on the location, objectives, research questions, research methods and the actual research instruments to be used. The research instruments also included the consent forms to be signed by participants. The researcher undertook to carry out the study in strict accordance with the approved research proposal and the ethics policy of Unisa. The ethical clearance approval letter is provided as Appendix A-1.

The researcher also obtained authorisation from the mine to conduct the research (Appendix A-2). A clear explanation of the research purpose and procedure was provided to participants prior to evaluations. Participants were asked to sign informed consent. As revealing their survey responses would not injure them in any way, it was decided not to use anonymity but rather confidentiality. Participants were ensured that, even though the findings of the evaluation would be used for research purposes and that the findings might be published in academic publications, their privacy would be protected by non-disclosure of their names, positions or affiliations. The informed consent document is given in Appendix B-2 as part of the user satisfaction questionnaire document.

For the heuristic evaluation, the expert evaluators were requested to sign consent forms. In this document, the evaluators were assured of anonymity, that their participation was voluntary, and that their inputs would be used purely for academic reasons. This informed consent form was part of the heuristic evaluation instrument, which is attached as Appendix B-1.

1.11. Structure of the thesis

The contents of this thesis cover the following areas:

- **Chapter Two: Virtual Reality**

This chapter sets the context of this study by describing the application domain of the research, namely virtual reality. It reports on a study of literature on VR and its applications, with specific reference to the mining industry.

- **Chapter Three: Systems Design, Usability and Instructional Design applied to E-learning Environments**

This chapter discusses the literature on human computer interaction (HCI) aspects relevant to this research, namely systems design, usability and instructional design. As this study focuses on VR training applications, learning theories are also discussed. The last section of this chapter discusses various usability evaluation methods.

- **Chapter Four: Mine Safety Practice**

Chapter Four focuses on current safety practices in the South African mining industry by introducing the South African mining environment and the major stakeholders involved in the industry, providing an overview of the safety legislation applicable to the mining industry, and presenting mine safety statistics.

- **Chapter Five: Research Design and Methodology**

Chapter Five explains the foundations and processes of this research. This chapter defines the underlying research paradigm of this study, the research and data collection methods used and the research design of the study. Furthermore, a heuristic framework comprising categories and criteria for the evaluation of desktop VR training systems is proposed.

- **Chapter Six: Usability Context Analysis**

This chapter discusses contextual analysis for the development of virtual reality applications, applied to safety training in mines. The results of the context analysis were applied to the design of a prototype, which was used and evaluated at a large platinum mine.

- **Chapter Seven: Prototype Design**

Chapter Seven presents the results of a case study relating to falls of ground at a large platinum mine, which led to the design of the second VR prototype. A design lifecycle model, synthesised for the development of VR training systems, is explained, followed by information on the detailed design and development of the second prototype.

- **Chapter Eight: Evaluation**

The empirical research described in this chapter relates to the application of the evaluation framework proposed by this study to evaluate the two VR prototypes, as well as the evaluation of user satisfaction.

- **Chapter Nine: Revised Evaluation Framework**

The outcome of the evaluations described in Chapter Eight not only provided valuable information regarding the prototypes, but also indicated that the evaluation framework itself had inadequacies. For this reason, a meta-evaluation of the Framework was done to strengthen the framework. The meta-evaluation and its findings are discussed in this chapter, followed by discussion of refinements to the evaluation framework, after which the improved evaluation framework is presented.

- **Chapter 10: Conclusion and Recommendations**

This chapter revisits the research questions and summarises the findings for each question; details the practical and theoretical contributions of the study; reflects on the implementation of the research design; discusses how the study implemented validity, reliability and triangulation; explains the limitations of the study; and makes recommendations and explores future work related to the research.

These chapters are followed by the list of references, as well as the appendices. Figure 1.3 presents the chapter layout of the thesis.

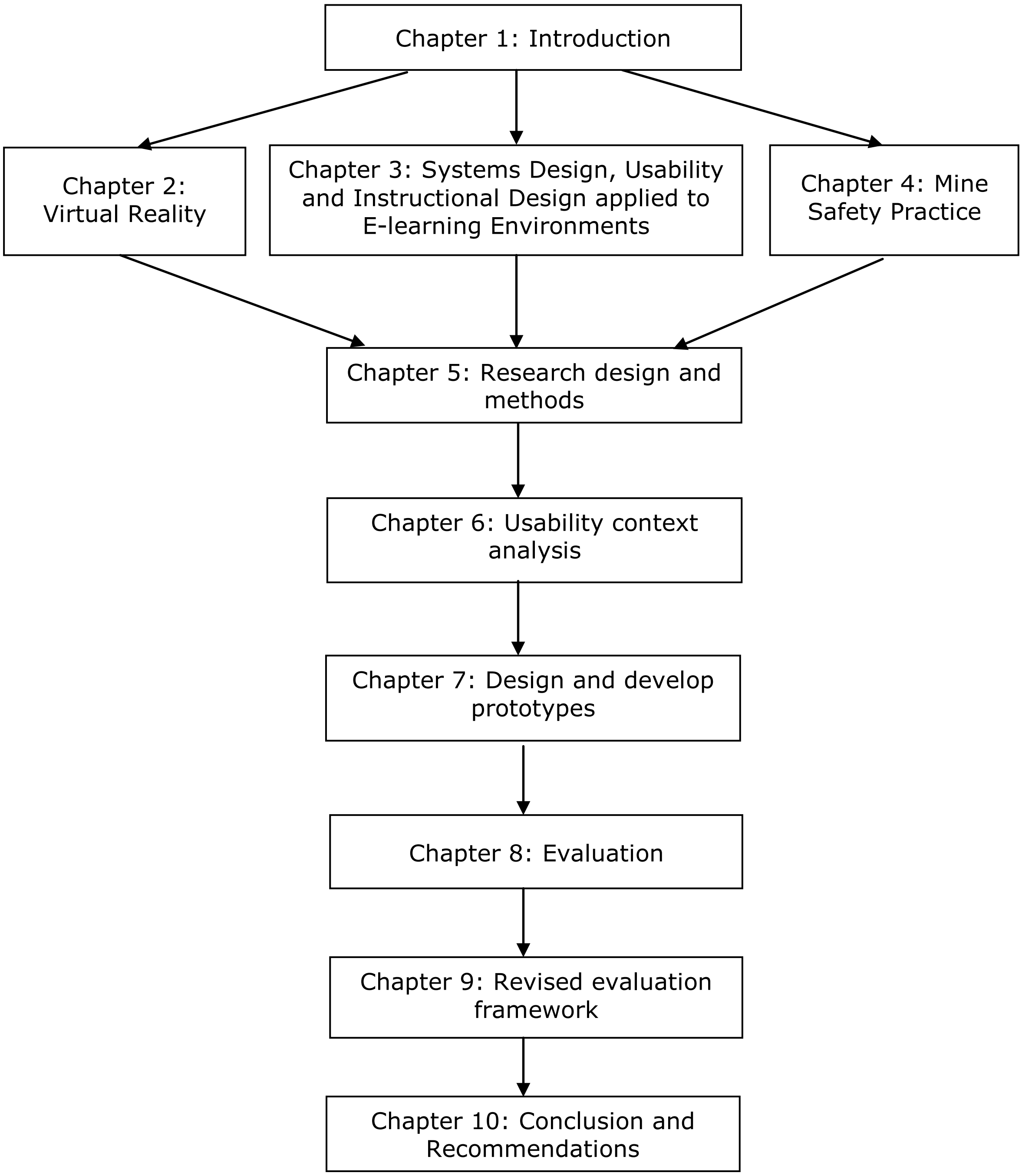


Figure 1.3: Structure of the thesis.

1.12. Chapter summary

This chapter introduced the topic of the research and discussed the background to the study. The problem statement and research questions were presented, and the rationale behind the study was described from mine safety training, educational and human–computer interaction perspectives. The value of the research was discussed, followed by a brief outline of the literature studies. The research design and methodology was explained, followed by a presentation of the research strategy. The scope of the study was discussed and the ethical considerations were explained. The last section provided an outline of the rest of this thesis.

The next chapter, Chapter Two, is a literature study on the application technology of this research, namely *virtual reality*.

Chapter Two
Virtual reality

2.1. Introduction

This chapter sets the context of this study by describing the application domain of the research, namely *virtual reality* (VR). It reports on a study of literature on VR and its applications, with specific reference to the mining industry.

In Section 2.2, various definitions of VR are presented and the categories and features of VR systems are discussed. The term *virtual reality* originally referred only to computer-generated virtual environments, but it is also frequently used in literature when referring to a mixture of real and virtual objects. A classification of available systems is presented to alleviate potential confusion when developing VR applications.

Section 2.3 of the chapter covers the application of virtual reality and several examples are presented where VR technology has been developed for research and industrial applications. The use of virtual reality for training purposes is covered in detail in Section 2.4 and, in particular, a report is given of the applications of VR for training in the mining industry in Section 2.5. This chapter addresses Research Subquestion 1 of this study: "What is the suitability and potential of virtual reality technology for training applications in the domain of mine safety training?" in Sections 2.4 and 2.5.

Figure 2.1 shows the chapter layout.

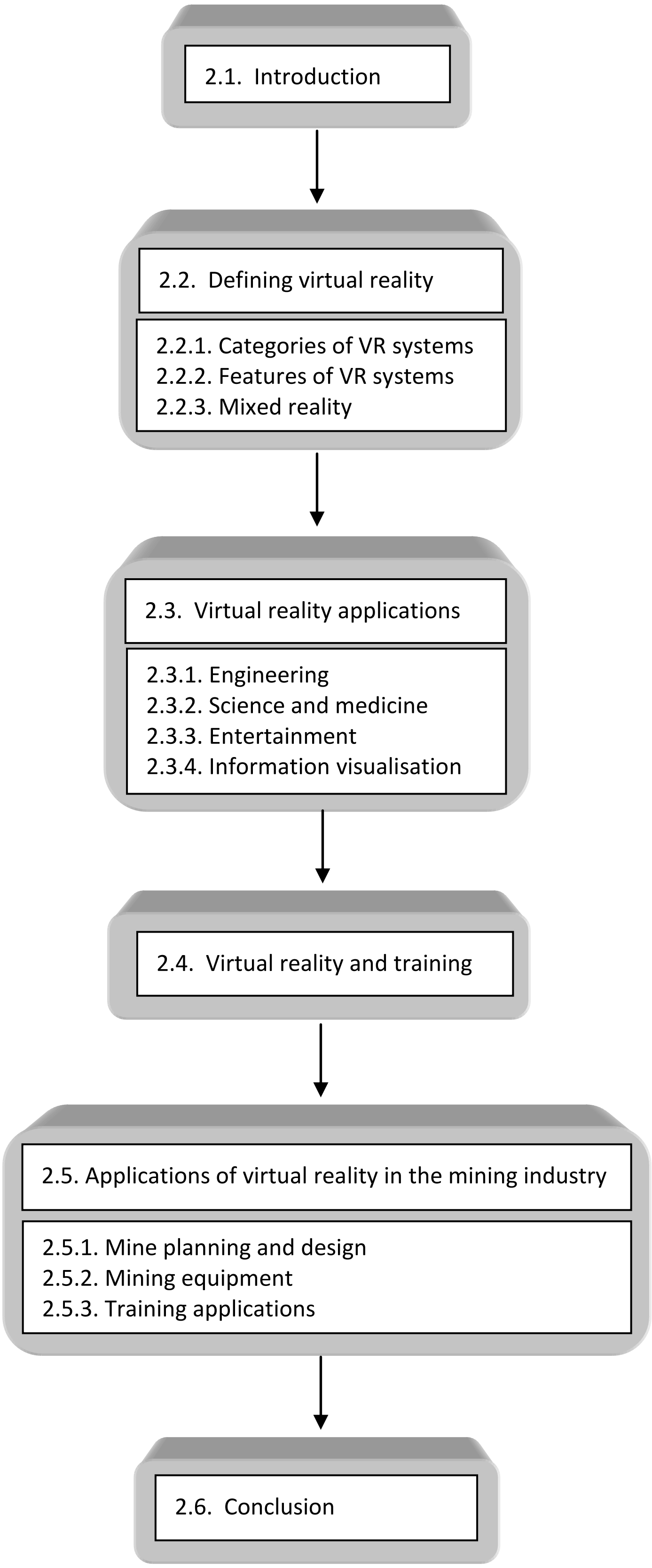


Figure 2.1: Layout of Chapter Two.

2.2. Defining virtual reality

Virtual reality technology is a computer-based technology that has developed rapidly since the 1990s. As computer hardware and software technology have improved, the ease with which interactive simulations can be developed and deployed has improved significantly and lower-cost, high-quality development tools have become available.

The term *virtual reality* was first used by Jaron Lanier to describe the immersive digital worlds he was trying to create (Rheingold, 1991). It is notable that there is still no generally accepted definition for virtual reality, but there is a plethora of visualisation technologies and computer-based visual content, all of which are commonly referred to as virtual reality. Various definitions and descriptions as to what constitutes VR can be found in the literature. A representative selection of such definitions follows, starting with some early definitions:

"Virtual reality can be described as the science of integrating man with information. It consists of three-dimensional, interactive, computer-generated environments. These environments can be models of real or imaginary worlds" (Roberts & Warwick, 1993:3).

"…virtual reality refers to an immersive, interactive experience generated by a computer" (Pimentel & Teixeira, 1993:11).

"Virtual reality is a synthetic, three-dimensional, interactive environment typically generated by a computer" (Iovine, 1995:2).

"…virtual reality means a computer-generated, interactive, three-dimensional environment in which a person is immersed" (Levy & Bjelland, 1995:xix).

The above definitions highlight four important aspects of VR namely, it is three-dimensional, interactive, immersive and computer-generated. The more recent definitions expand on the facets of VR and elaborate on the nature of current virtual environments, indicating that VR is an evolving, cutting-edge technology. Some examples are:

"Virtual reality (VR) is a technology that allows a user to interact with a purely computer-simulated environment. The simulated environment can be similar to the real world, for example, simulations for aircraft pilot or combat training, or it can differ significantly from reality, as in computer games. Virtual reality is limited only by imagination" (Stothard, Squelch, Van Wyk, Schofield, Fowle, Caris, Kizil & Schmid, 2008:8).

"VR is a continuously evolving technology that provides three-dimensional spatial environments through advanced forms of computer graphics" (Poynton, 2009:8).

"VR is described as a cutting-edge technology that allows learners to step through the computer screen into a three-dimensional interactive environment" (Chen, 2010:13).

Kalawsky (1997) went beyond a definition and gave a broad description of what a VR system allows a participant to do, namely to:

- become immersed in a completely synthetic computer-generated environment;
- achieve a sense of presence in the environment;
- become uninhibited where conventional laws of physics can be controlled in a way that assists greater understanding;
- achieve a sense of non-real time, where situations can be presented in slow or fast time;
- achieve a high degree of interaction that can equal or exceed that achievable in the real world;
- interact in a completely natural and intuitive manner with the synthetic environment;
- repeat the task until the desired level of proficiency or skill has been achieved; and
- perform in a safe environment.

All the aspects of Kalawsky's structure of elaboration are equally valid today.

From the above definitions and descriptions it is clear that VR refers to a technology where a user interacts with a three-dimensional computer-simulated environment, which the user perceives as comparable to real-world objects and events. Such environments are frequently referred to as virtual environments. The aspect of immersion is also evident in some definitions, but full immersion is not necessarily considered essential in all applications.

2.2.1. Categories of virtual reality systems

Moreno and Mayer (2002) identify three categories of VR systems, namely fully-immersive, semi-immersive, and desktop or non-immersive systems. Immersion is the extent to which the senses of a user perceive the illusion of reality delivered by the VR system's display. The concepts within Moreno and Mayer's categories are used as subheadings in this section.

2.2.1.1. Immersive virtual reality systems

Immersive systems are the most technically advanced applications of virtual reality. The user is essentially isolated from the outside world and fully enveloped within the computer-generated environment. The user's view of the real world is replaced by computer-generated images that react to the position and orientation of the user's head.

The user in an immersive virtual reality system is required to wear a head-mounted display (HMD) that presents an image directly in front of each eye and magnifies it so that it fills a wide field of view, creating the impression of actually being within an environment, rather than gazing at a screen. As the user looks around, the position and orientation information is continuously relayed to the host computer. The computer calculates the appropriate view (virtual camera view) that the user should see in the virtual environment, and this is displayed on the miniature displays in the HMD (Daden, 2014).

Objects in the virtual world may be manipulated by means of a data glove. A data glove measures the flexure (bend) of the user's fingers. The user may grab a virtual object and move it to a different spot. The position and orientation of the user's hand is measured with a Six Degrees of Freedom tracker (5DT, 2012). For example, if a user turns to look backwards over the left shoulder, a sensor will detect the change in position and orientation of the head, and adjust the visual display so that the display corresponds to what the user would see from that position if the scene were real. If the user reaches out towards an object in the virtual environment, sensors sensitive to movements of the fingers and to the position of the hand, enable the system to detect when the user's hand intersects with the virtual object, and adjust the display to mimic pushing, lifting, or rotation of the object (Weiss & Jessel, 1998). Example photographs of a head-mounted display and a data glove are shown in Figure 2.2.

Head-mounted display technology has a number of disadvantages including encumbrance, a sense of isolation resulting from the experience, high cost, and occasional simulator sickness and disorientation (Kalawsky, 2006; Rogers, Sharp & Preece, 2011; Stone & Knight, 2012).

Figure 2.2: An example of a head-mounted display and a data glove (5DT, 2012).

Oculus VR is currently developing a lightweight, low-cost, stereoscopic 3D headset, called the *Oculus Rift*. The *Oculus Rift* has a 100° field of view and uses tracking technology to provide ultra-low latency 360° head tracking, allowing the user to seamlessly look around in the virtual world (Oculus, 2015). *Facebook* has recently invested $2 billion in the research and development of the *Oculus Rift*, while *Sony* is also developing a similar product, called the *Sony Morpheus HMD*. It is expected that the affordability of these products may lead to mass-market applications of VR, especially in the domain of computer gaming (Daden, 2014).

To date only the development kit of the *Oculus Rift* has been released and the commercial product is not yet available. Nevertheless, some research studies on its application have already been published, for example, on using the *Oculus Rift* for visualisation of complicated biological or molecular structures (Lartigue, Scoville & Pham, 2014) and utilising the *Oculus Rift* for high-definition, immersive, distance learning (Lartigue, Cathcart, Kelleher, Pfundstein & Williams, 2014).

2.2.1.2. Semi-immersive virtual reality systems

Semi-immersive VR systems are also called Projection VR (Nasios, 2001; Squelch, 2001) or Multi-display VR (5DT, 2012). In semi-immersive VR systems, computer-generated images are displayed on large screens by a stereo projection system and are viewed via special stereo eyewear. Interaction with onscreen menus is achieved via a remote keypad and other input is handled by devices such as 3D controllers or joysticks.

The use of multiple projection-based systems can result in significant cost, but high resolution images can be produced. The field of view (FOV) of the user is extended by using several computer monitors, or projectors. When using projectors, the image may be front-projected or back-projected onto the viewing screen. Many simulators utilise three screens to provide an extended FOV (forward view, left view and right view). The configuration in which the user is surrounded by projection screens, is sometimes referred to as a CAVE (Cave Automatic Virtual Environment). A CAVE creates the illusion of immersion by projecting high-resolution stereo images on the walls and floor of a room-sized cube. Several persons wearing lightweight stereo glasses can enter and walk freely inside the CAVE (Sveistrup, 2004).

An example of a 360° stereoscopic system is the AVIE (Advanced Visualisation and Interaction Environment), developed at the University of New South Wales in Australia (Shaw & Del Favero, 2004). AVIE uses motion and shape tracking systems and a multi-channel audio system that enables the development of immersive visualisation applications. Three-dimensional audio-visual experiences are created by a combination of multiple projectors and multiple sound sources.

Users within the visualisation environment are tracked by a system of infra-red cameras and real-time software able to generate models of their movements and body positions. These features enable audience participation, interplay between real people and projected characters or avatars, and analysis of trainee behaviour.

In projected VR systems the images may also be projected on a dome that may vary in shape and size. An example is iDome, which is configured as a fibreglass hemisphere with a three-to-five metre diameter that stands vertically in front of the viewer, with a projector, computer, surround audio equipment and user interface (UNSW, 2013). The AVIE and iDome are shown in Figures 2.3 and 2.4 respectively.

Figure 2.3: The AVIE (Advanced Visualisation and Interaction Environment) developed by the University of New South Wales (UNSW, 2013:2).

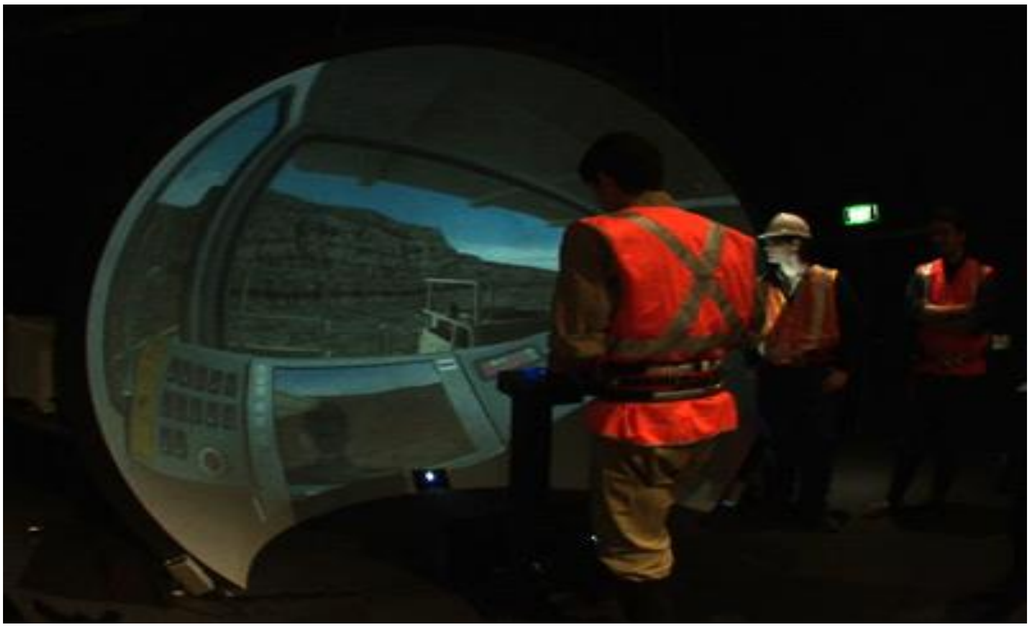


Figure 2.4: The iDome hemisphere (UNSW, 2013:5).

Although their fixed display position and limited display area restrict the user's range of interactions, projected VR systems have an advantage in that the user is not tethered by a glove and headset and is able to communicate freely with non-VR participants.

2.2.1.3. Desktop virtual reality systems

Since not all applications require immersion to the extent presented in the previous two categories, more affordable, non-immersive VR systems provide practical alternatives. Desktop virtual reality systems are the most popular type of virtual reality systems and are based on the concept that the user interacts with the computer screen without being fully immersed and surrounded by a computer-generated environment. The user remains visually aware of the real world, but is also able to observe the virtual world on a high-resolution monitor.

These systems provide a lower level of presence than immersive systems, but the lower capital cost of hardware, software and peripherals make desktop virtual reality systems an attractive and realistic alternative, available at a reasonable price to many end-users (Nasios, 2001). Desktop systems utilise standard computer hardware, and input devices include a keyboard, mouse, 3D controller, joystick, trackball, force ball or voice to interact with and manipulate the virtual environment. The sense of subjective immersion in desktop virtual reality systems can be improved through stereoscopic glasses, which give the extra dimension of three-dimensional depth.

Desktop virtual reality systems provide a low-cost option for high-resolution visualisation for design, training and education applications.

2.2.1.4. Discussion

The boundaries between the different categories of VR systems are not clear-cut, since the creative use of display and auditory peripherals in desktop or semi-immersive systems can promote a sense of presence as experienced in immersive systems, even in the absence of the ability to fully control the virtual environment.

In networked VR, it is also possible for different users to share the same virtual world. This is achieved by connecting the host computers to a computer network. Each user's host computer broadcasts the position and orientation of the user in the virtual world. The users can then see representations of each other in the virtual world and they can interact, work together or compete (5DT, 2012).

2.2.2. Features of virtual reality systems

From the definitions of VR, a number of key features of VR applications can be identified. These include *computer-generated, three-dimensional* and *interactive*. Other cardinal features discussed in literature are *presence, viewpoints, realism* and *fidelity*. These features are discussed in the following subsections.

2.2.2.1. Computer-generated and three-dimensional

The term *computer-generated* means that the virtual environment is created on, and rendered and supported by, a computer. Three-dimensional (3D) refers to the inclusion of a third dimension, that is the **z** dimension over and above the existing **x** and **y** dimensions. This makes it possible to simulate depth of objects 'into' the computer screen, enabling the description of space and full rotation of virtual objects.

The process of creating 3D computer graphics involves 3D modelling, animation and rendering. An on-screen 3D simulation of an object (real or imagined) is called a 3D model. In order to make the model more realistic, visual characteristics such as shading, shadows and textures can be added, and movements can be simulated through computer animation. Rendering is the process of calculating the appearance of the 3D model, which enables an entity to be drawn on a two-dimensional screen, yet appear to be in three dimensions.

2.2.2.2. Interaction

Earlier definitions of interaction in VR research related to the modification of, and navigation in, virtual environments (Zeltzer, 1992; Steuer, 1992; Slater & Usoh, 1994). Heeter (2000) defined interaction as a series of actions and reactions of a user with the virtual environment. The predominant definition in VR research today describes interaction as the ability of the user to move within the virtual world and to interact with the objects of the virtual world (Daden, 2014; Nalbant & Bostan, 2006). This includes the ability of the user to select objects displayed in the virtual environment, manipulate them and acquire information about them.

Virtual reality systems provide greater levels of interactivity than other computer-based systems. This can be very beneficial in training environments, provided that the interfaces are intuitive and easy to use (Kalawsky, 2000). VR systems allow the user to interact not only with the virtual environment, but also with virtual objects inside the environment. In desktop VR systems, the interaction is usually achieved via a mouse, joystick, keyboard controls or touch screen. In more immersive systems, movement of a 'virtual hand' allows a more natural interaction with objects.

2.2.2.3. Presence

An important feature of virtual reality is the provision of a sense of actual presence in the simulated environment. Presence refers to the subjective experience of 'being' in the computer-generated environment, rather than in the actual real-world environment (Slater & Wilbur, 1997). It is related to immersion, where virtual environments with higher levels of immersion induce a higher sense of presence (Welch, Blackman, Liu, Mellers & Stark, 1996).

A number of factors have been identified that account for the strength of presence. These are generally categorised into control factors, extent of sensory information, distraction factors, ease of navigation and realism factors (Lewis & Griffin, 1997; Witmer & Singer, 1998). Achieving presence in applications of virtual reality depends on the goals of the particular application, as well as on the cost and technical complexity that its developers are willing and able to assume. One of the most important consequences of presence in VR systems is that a virtual experience can evoke the same reactions and emotions as an experience in the real world (Rogers *et al.*, 2011; Van der Straaten, 2002).

2.2.2.4. Viewpoints

Virtual reality offers first-person and third-person viewpoints to the user. In a first-person perspective, users view the virtual environment through their own eyes, whereas in a third-person perspective, the virtual environment is viewed through a character visually represented on the screen. These characters are called avatars and could be a representation of users themselves or a predefined game character as occurs in 3D computer games.

An example of a first-person perspective is what is experienced in first-person shooter games. The player moves through the virtual environment without seeing a representation of him/herself. Usually the barrel part of a gun is shown to indicate where the player's gun is aimed. As the player moves around in the environment, the barrel of the gun indicates the player's current position. In a third-person game the player sees the avatar as well as the virtual world surrounding him/her. The player then controls the avatar's interactions with the environment by controlling the avatar's movements.

In VR applications, first-person perspectives are typically used in simulators, where it is important to have direct or immediate control over interaction mechanisms, such as steering wheels or equipment controls. Third-person perspectives are used in learning environments or simulations where it is important for the users to see themselves in relation to others or objects in the environment. In some virtual reality applications it is possible to switch between the first-person and third-person perspectives, enabling the user to experience different viewpoints on the same virtual environment (Rogers *et al.*, 2011).

2.2.2.5. Realism and fidelity

One of the major research directions in computer graphics is realistic image synthesis. The pursuit of realism has led to many major breakthroughs in 3D modelling, display algorithms and rendering. Producing realistic 3D images is both an art and a science and is expensive time-wise and computationally, causing some researchers to question the need for and the value of close realism in computer graphics applications (Ferwerda, 2003; Gershon, Braham, Fracchia, Glassner, Mones-Hattal & Rose, 1996; Strothotte & Schlectweg, 2002).

Ferwerda (2003) distinguishes between three varieties of realism when producing an image of an actual scene.

- Physical realism: The image provides the same visual *stimulation* as the scene. This means that the computer-generated model accurately reflects the shapes, materials and lighting properties of the scene.
- Photo-realism: The image produces the same visual *response* as the scene, so that it appears as though it was produced by photographing the scene.
- Functional realism: The image provides the same visual *information* as the scene. 'Information' here refers to properties of objects in a scene, such as sizes, shapes, motions and positions. Functional realism is defined in terms of the *fidelity* of the information provided by the image. An image is functionally realistic if it allows the observer to make reliable visual judgments and to perform useful visual tasks.

In the context of VR, fidelity is defined as the degree to which the virtual environment emulates the real world. Poynton (2009) defines three categories of fidelity.

- Physical fidelity: The degree to which the virtual environment looks, sounds and feels like the actual environment.
- Functional fidelity: The degree to which the virtual environment behaves like the actual environment when responding to user interaction.
- Psychological fidelity: The degree to which the virtual environment replicates the psychological factors experienced in the real-world environment, for example, stress or fear.

Virtual reality systems should provide adequate levels of realism. When looking at the wide variety of technologies available within a considerable range in price, VR systems developers need to know how realistic virtual environments have to be in each particular application, and to what degree a sense of presence is required, in order to accomplish the objectives of the virtual experience. In the VR community, it is often assumed that the higher the level of realism the better, but certain researchers (Dennis & Harris, 1998; Goldstone & Son, 2005; Jentsch & Bowers, 1998; Smith, 2003) point out that VR environments may be most effective when only the appropriate details are embellished. This increases realism, even when the fidelity deviates noticeably from the real world.

An assessment should be made to determine whether the benefits of highly detailed models justify the extra time, cost and effort taken in building them. The tasks a VR user needs to perform will determine the kinds of visual information that should be faithfully represented by the computer-generated images. High realism may be necessary in applications such as medicine, aviation or mining, which require precise rendering to simulate real, life-threatening situations. In other application areas, such as military systems, it may not be necessary for a virtual world to be an exact depiction of the real world. In such application areas, it is sufficient to have a virtual world in which it is possible to work in much the same way as in the real world (NATO, 2003).

In the case of e-training, VR visual content must not only be credible and believable to trainees (*physical fidelity)*, but it must also foster an effective assimilation of the training goals and be capable of supporting the transfer of skills and knowledge learned to real-world settings, minimising skill or knowledge fade over time (*psychological fidelity*) (Stone, 2012).

2.2.3. Mixed reality

In the conventional view of a VR environment, the user is to some degree immersed in, and able to interact with, a completely synthetic world. The VR label is also frequently used, however, in association with a variety of other environments where a mixture of real and virtual objects may be used. In a landmark paper published in 1994, Milgram and Kishino introduced the concept of a *virtuality continuum* which relates to the mixture of objects presented in any particular display situation, where *real environments* are shown at one end of the continuum and *virtual environments* at the opposite end (Milgram & Kishino, 1994).

The merging of real and virtual worlds is generically referred to as *Mixed Reality (MR).* Mixed reality is a term that is all encompassing and spans much of the virtual continuum, shown in Figure 2.5. All real and synthetic aspects of real and virtual environments fall somewhere along the virtual continuum and many of the terms presented by Milgram and Kishino (1994) are currently in regular use.

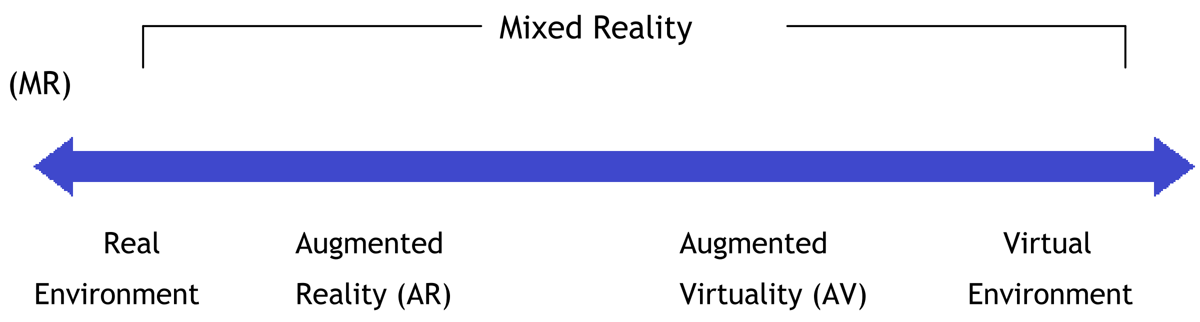


Figure 2.5: Simplified representation of the virtual continuum (Stothard *et al.*, 2008:11, adapted from Milgram & Kishino, 1994).

The term *reality*, in its widest sense, includes everything that is, whether or not it is observable or comprehensible. Augmented reality (AR) refers to the overlaying of computer-generated imagery or data onto real-world imagery. AR technology makes it possible to project data, diagrams, animation or video onto transparent glasses, which the user can then see while viewing the real world (Van Krevelen & Poelman, 2010). AR research focuses on the use of video imagery which is digitally processed and 'augmented' by the addition of computer-generated graphics.

Augmented virtuality (AV) refers to the merging of real-world objects into virtual worlds. Physical elements, such as physical objects or people, are dynamically integrated into, and can interact with, the virtual world in real time.

Interactive visualisation technology is advancing rapidly and is becoming more widely available to potential users. In general, increased accessibility has reduced the cost of VR hardware and software. The spectrum of technology ranging from hi-end proprietary systems through to off-the-shelf game-based systems has made entry easier for developers and users alike. The options available to developers new to the concept can be confusing. Determining which system has what functionality and what content can be displayed by each system, is also difficult. A classification of available systems assists in alleviating this confusion.

Milgram and Kishino (1994) formulated a taxonomy of the various ways in which the 'virtual' and 'real' aspects of MR environments can be realised. The taxonomy distinguished between six classes of visual displays. This classification was first published in 1994 and, since then, the application of head-mounted displays has not been as prominent as was predicted. In fact, other technologies have emerged and become popular, such as mobile devices, domes and CAVE systems. In an attempt to quantify visualisation systems, their uses and their implementation, Stothard *et al.* (2008) expanded this taxonomy to include ten classes of visual displays.

- Class 1: Monitor-based (non-immersive) video displays – that is, conventional computer-screen displays upon which computer-generated images are electronically or digitally overlaid. This includes technology such as chroma-keying and stereoscopy.

- Class 2: Video displays as in Class 1, but using immersive HMDs, rather than conventional monitors.

- Class 3: HMDs equipped with a see-through capability, with which computer-generated graphics can be optically superimposed onto directly-viewed real-world scenes.

- Class 4: Same as Class 3, but using video, rather than optical viewing of the 'outside' world. The difference between Classes 2 and 4 is that with Class 4 the displayed world should correspond orthoscopically with the immediate outside real world, thereby creating a 'video see-through' system, analogous to the see-through of Class 3.

- Class 5: Completely graphic display environments, completely immersive, semi-immersive or otherwise, to which video 'reality' is added.

- Class 6: Completely graphic but semi-immersive environments (e.g. large screen displays) in which real physical objects in the user's environment play a role in the computer-generated scene, such as in reaching in and ‘grabbing’ something with one's own hand.

- Class 7: Large full-surround screen(s), completely graphic environments that are fully immersive and use real physical objects or haptic devices to play a role in the computer-generated scene. Examples of these are CAVE and visionarium systems and, in some cases, domes.

- Class 8: Devices with a capability to see through to the real world and simultaneously show computer-generated graphics.

- Class 9: Hand-held mobile devices with the capability to show video of real-world and computer-generated graphics simultaneously.

- Class 10: True 3D holographic or hologrammatic representations that mix reality with the real world.

The taxonomy provides insight into where technology can be implemented in the development of virtual environments. The classification and the application of each class are shown in Table 2.1.

Table 2.1: Taxonomy of Interactive Computer-based Visualisation Systems, based on Stothard et al. (2008).

Type	Class Description	Immersive – Full, Semi or Non	Individual or group use
Class 1	Monitor-based video displays with capability to show video and computer-generated graphics simultaneously	Non	Individual
Class 2	Head-mounted video displays with capability to show video and computer-generated graphics simultaneously	Full	Individual
Class 3	Head-mounted video displays with capability to see through to real world and show computer-generated graphics simultaneously	Full	Individual
Class 4	Head-mounted displays with capability to show video of real world and computer-generated graphics simultaneously	Full	Individual
Class 5	Large screen(s), completely graphic display environments with capability to show computer-generated graphics and video simultaneously	Full & Semi	Individual or group
Class 6	Large screen(s), completely graphic environments that are partially immersive and use real physical objects or haptic devices to play a role in the computer-generated scene	Semi	Individual or group
Class 7	Large full-surround screen(s), completely graphic environments that are fully immersive and use real physical objects or haptic devices to play a role in the computer-generated scene	Full	Individual or group
Class 8	Device with capability to see through to real world and simultaneously show computer-generated graphics	Non	Individual or very small group (2-3 people)
Class 9	Hand-held mobile device with capability to show video of real world and computer-generated graphics simultaneously	Non	Individual or very small group (2-3 people)
Class 10	True holographic devices	Semi	Individual or group

2.3. Virtual reality applications

Apart from many VR training applications, other applications being developed in the field of virtual reality span a wide spectrum of domains, from games to construction and from visualisation to technology demonstrators.

2.3.1. Engineering

VR offers new possibilities in the field of product development by speeding up the pace and improving the quality of designs (Ottosson, 2002). Through virtual prototyping, VR technology is used to virtually test prototypes of products before final verification with physical prototypes is performed, which can lead to improved ergonomic design and usability. In this way, customers and users can interactively influence the features of a product idea before the product actually exists.

Engineering applications using VR technology include focus areas such as computer-aided design and manufacturing of products and systems, architectural design and ergonomic issues. VR technology can generate virtual engineering spaces where objects can be manufactured, inspected, assembled, tested, and subjected to a range of simulations. An example of a situation where VR is applied effectively as a tool for product design and design review, is the automotive industry. Virtual prototyping is an effective means of shortening development times for new cars. Most major automobile manufacturers have VR installations, for instance, *Daimler-Chrysler's* virtual ergonomics facility and the *Rover* group's virtual engine assembly facility (Joung & Noh, 2014).

Other companies where VR is used in design reviews include the Electric Boat Division of *General Dynamics Corporation*, the company that is the primary builder of nuclear submarines for the United States Navy (VRSim, 2015), as well as *John Deere*, the world's largest company in farm and forestry equipment (John Deere, 2013).

The *Electric Power Research Institute* (EPRI, 2003) reports on the use of VR in nuclear plants. In what is called virtual reality construction, 3D plant models are used to evaluate, debug and refine the design and construction of plant components. This has reduced costs in the construction of nuclear plants, optimisation of construction sequencing and labour deployment, identification and elimination of problems before they were encountered in the field, and improved preparation of labour crews.

Noor and Wasfy (2001) used VR to simulate physical experiments concerning wind-tunnel testing and structural testing for aerospace systems. They created an immersive virtual environment comprising rear-projected surround screens, head-mounted tracking devices and voice-recognition input capabilities. A number of virtual laboratories were created within which users could interactively conduct experiments in a risk-free environment.

The following are further examples of VR applications in the engineering environment.

- Experimental avionics test beds: New avionics configurations and layouts can be evaluated while using virtual aircraft to fly over landscapes that closely resemble reality.
- Marine technology demonstrators: Ships and submarines are modelled and simulated.
- Virtual battlefields: Various entities in the battlefield can be controlled by human operators and realism can be enhanced by realistic sound effects, explosions, fire, smoke and dust (Werning, 2009).

2.3.2. Science and medicine

The use of virtual reality for interpretation of scientific datasets is well established in fields such as molecular modelling, cartography, archaeology, oceanography and medicine. Using powerful computer graphics workstations, images in these domains can be rendered in real time and parameters within the simulation exercise can be adjusted with immediate responses. These techniques offer an effective problem-solving environment and provide many benefits to scientists.

The following are specific examples of VR applications in these fields.

- Virtual reality-based balance exercise programmes for adults with traumatic brain injury (Thornton, Marshall, McComas, Finestone, McCormick & Sveistrup, 2005) resulted in significant improvements being noted in balance confidence and function of the participants.
- Psychiatric treatment at *Georgia Institute of Technology* and the *Emory University Medical School* uses VR simulation to treat fear of flying, fear of heights and fear of public speaking (Rothbaum, Anderson, Zimand, Hodges, Lang & Wilson, 2006).
- At the *Atlanta Veterans Administration Hospital,* war veterans are treated for post-traumatic stress disorder. Patients are led into a simulated battle scene, recreating a situation in which they can relive their stress experiences. By leading the patient completely through the scene and out the other side, the psychologist

aims to help the patient to move beyond the damaging patterns (Ready, Pollack, Rothbaum & Alarcon, 2005).

Another interesting example relates to VR pain distraction systems, used to distract the attention of patients who undergo painful procedures, especially where anaesthesia is not a viable option. Because the patient wears a head-mounted display, he/she is fully immersed and cannot see the procedure being performed. Games can also be used as distraction to younger patients (5DT, 2012).

Virtual reality is also being applied in the fields of nanotechnology and bionanotechnology (Sharma, Mavroidis & Ferreira, 2005). VR interfaces to real-time simulations hold great potential for molecular and nanotechnology scientists, since immersive visualisation in three dimensions allows the scientist to gain a deeper understanding of the micro- or nanoworld.

2.3.3. Entertainment

In the early 1990s, a number of new software development businesses started to make VR systems for games in arcades, but the computers of the time could provide only primitive graphics. Furthermore, the head-mounted displays and gloves were uncomfortable. Many of these companies went bankrupt as disappointed players stopped purchasing such games (Scienceclarified, 2009). However, current arcade games are once again including aspects of virtual reality. Examples include virtual bowling, where players can roll virtual bowling balls through three-dimensional landscapes displayed on large screens, or strap-in simulators where users can experience various simulated rides, from virtual roller-coasters to flights through space.

Fantasy role-playing games are developed for the Internet, where thousands of players can play a game online simultaneously, interacting with one another. Initially these online multi-player environments contained only text, but due to advances in computer graphics technology, online games can now contain virtual worlds utilising high-resolution graphics.

Further examples of online environments are *ActiveWorlds* and *Second Life*, where members of virtual communities can build homes on virtual ground, set up businesses, play games, shop, fly or chat with one another. Members join the virtual community online and must pay a monthly subscription fee. In many cases, their activities in the online community are limited only by their imaginations (ActiveWorlds, 2012; Second

Life, 2012). In such virtual communities, people represent themselves in the virtual environment as personas called avatars. The meaning of avatar is a concept similar to that of incarnation, which means 'taking on flesh'. Avatars in 3D computer games are essentially the players' physical representation in the game world. In most games, the player's representation is fixed, but games increasingly offer a basic character model, or template, and then allow customisation of the physical features as the player sees fit.

Computer-generated imagery is increasingly used in movies to create special effects. As an example, in the science fiction movie *Avatar,* innovative digital 3D techniques were used to create the movie, with cameras being specially designed for the film's production. The movie combines real photography and computer-generated imagery. A specially-designed camera was built into a boom that allowed the facial expressions of the actors to be captured and digitally recorded for the animators to use later. The stereo in the movie was created by rendering images of two cameras used side by side (Joubert, 2009).

A further example of the use of VR in entertainment is *Disney's* indoor interactive theme park called *DisneyQuest*, where visitors can experience various virtual reality adventures, including a virtual river-raft cruise through a jungle in the *Dinosaur Age*, a chance to pilot *Aladdin's* magic carpet, and a battle against villains in a virtual reality comic book world (Disney World, 2013).

2.3.4. Information visualisation

Computer-generated visualisations of complex data enable users to see patterns, trends and anomalies that enhance discovery, decision-making and explanation of phenomena. One of the techniques that is used for depicting information and data is the generation of 3D interactive maps, which present data via trees, clusters, scatter plot diagrams and interconnected nodes (Chen, 2004).

Sarathay, Shujace and Cannon (2000) developed a prototype VR-based visualisation tool for the visualisation of multi-gigabyte datasets. This tool used virtual reality to display information within a virtual environment that allows non-expert users to examine, comprehend and interpret the data. User navigation was achieved via a simple fly-through method whereby users could explore the dataset internally with no restrictions on movement through layers of data.

Ziegeler, Moorhead, Croft and Lu (2001) investigated the visualisation of meteorological data in a 3D immersive virtual environment. Weather information is traditionally displayed as two-dimensional plots and it is difficult for forecasters to view the entire picture of the atmosphere. With multi-layer, time-series data it is difficult to see all layers and time steps in a single image. Ziegeler *et al*. used animation in an immersive environment to successfully fuse multiple layers, time steps and variables into a 3D visualisation of the data. This method can lead to improved interpretation of meteorological data.

VR is also used to generate virtual landscapes closely resembling real landscapes. This is done by integrating geographical information system data, satellite photographs and aerial photographs. These landscapes can be referenced to electronic maps of the area. Other examples of visualisation applications include:

- airspace visualisation, which enables air traffic controllers to visualise the airspace surrounding an airport,
- flight visualisation, where a pilot can navigate through the virtual landscape during mission planning, familiarising him/herself with key features of the landscape, and
- a virtual wind-tunnel, which uses computational fluid dynamics datasets to calculate windflow lines to allow flow visualisation (5DT, 2012).

Virtual reality and augmented reality technologies are used in museums, where virtual museums are constructed, and are accessible over the Internet or through kiosks located in accessible places within the museum (Liarokapis, Sylaiou, Basu, Mourkoussis, White & Lister, 2004; Manic, Aleksic & Tankosic, 2013; Misu, Georgila, Leuski & Traum, 2012; Preradović, Miličić & Duričić, 2014). To make the museum collections accessible to people with physical disabilities, user interactions in the virtual museum are performed with the help of assistive technology, so that users can benefit in terms of education and entertainment.

2.4. Virtual reality and training

The next two sections address Research Subquestion 1 of this study.

RQ1	What is the suitability and potential of virtual reality technology for training applications in the domain of mine safety training?

Virtual reality has also been a focus of research interest with regard to its potential applications in training (Kinshuk, Lin & Patel, 2008). Virtual reality-based training has been accepted in various industries as an effective method of training. VR training can be applied to a variety of workplace activities, including those of a safety-related nature (Squelch, 2001). The use of VR training tools can help to reduce accidents or incidents that cause injuries and fatalities, since they allow employees to practise skills from the safety of a computer-based simulated environment.

VR is ideal for the training of workers who perform tasks in dangerous or hazardous environments. The trainee can first practise the procedure in a risk-free virtual environment, and can be exposed to 'life-threatening' scenarios in a safe and controlled environment (STS, 2013). Other advantages of VR training are that VR can expose trainees to realistic, functional simulations, and trainees can demonstrate mastery of skills through performance of tasks in multiple scenarios (Rebelo & Noriega, 2012).

VR training can be used productively in various situations (Chen, 2010; Mitra & Saydam, 2013; Webber-Youngman, 2014; Weiss & Jessel, 1998):

- It is an ideal tool to train equipment operators. VR safeguards expensive equipment as the trainees practise in a simulated environment. This is also beneficial if the actual equipment has high running costs.
- It is appropriate for emergency scenarios and dangerous situations where real-world training is not feasible.
- VR has the ability to visually represent highly-technical content.
- Trainees can explore a wide variety of scenarios in a high-retention 3D virtual environment.
- The relevant virtual environment cannot be experienced in the real world.
- VR can be used to learn to perform routine tasks without pressure and to learn simple components of more complex tasks.
- VR has the ability to provide the trainee with viewpoints from different angles of the three-dimensional environment.

Some of the specific fields and disciplines where VR training is used, include astronaut training, education, the automotive, aeronautical and aviation industries, the military, electrical environments, marine and medical fields. Each of these will be discussed separately in the next subsections, followed by a separate section on the use of VR for simulation and training in the mining industry, which is the focus of this study.

2.4.1. Astronaut training

A number of VR training systems are used by the *National Aeronautics and Space Administration (NASA)* and several prototypes are also being developed for possible future use in a dedicated VR lab. The VR system that has been used the longest at *NASA* is the system used for training astronauts for extra-vehicular activity. Tasks such as correcting the optics of the *Hubble* telescope mirrors, caused new training demands. This VR environment allows astronauts to practise the careful planting of hands and feet while moving around on the outside of a space vehicle (Osterlund & Lawrence, 2012).

NASA's VR facility uses computer graphics displayed on head-mounted displays. As input devices, data gloves are used in combination with haptic feedback devices, which provide a 'touch' response to actions. This equipment simulates the look and feel of doing a spacewalk. The spacewalking astronauts can see all the space station or space shuttle structures, which are graphically presented in the VR headsets (Homan, 2010). An astronaut can move around the simulated space station by using data gloves to 'grab' handrails located on the structure.

The VR lab also provides flight simulator training for using the jet backpack worn by all spacewalkers while they are working outside the space station. The jet backpack unit provides a spacewalker with the capability to fly back to the space station should he or she become detached and float away from the station. The system uses a hand controller and a head-mounted display to enable the spacewalker to practise flying the unit.

2.4.2. Education

"Virtual reality is emerging as a very powerful educational tool that has the potential to provide education establishments with a powerful and effective educational environment" (Kalawsky, 2006:1). A major advantage of VR is the way it permits students to interact with educationally-orientated simulations. Virtual reality has become an important part of teaching and training, transforming the way people work and learn, because virtual environments improve the possibility of learner engagement (Burkle & Kinshuk, 2009). VR is listed as one of the top ten smart technologies for schools in the USA and VR-equipped classrooms are used to practise life skills, reading and mathematics (Warlick, 2002).

An example of virtual reality at secondary school level is a program called *Touch the Sky – Touch the Universe*. This system teaches students about astronomy by presenting an interactive 3D model of the solar system. Students can fly a virtual spaceship through the virtual environment to view planets, moons, comets and asteroids, to watch eclipses occur, and to observe planets move through their orbits (Yair, Mintz & Litvak, 2001). Another educational VR program teaches the development of modern architecture by allowing students to 'fly' over Chicago and explore more than forty of the city's buildings. The system also includes a demonstration of how a skyscraper is built (SunriseVR, 2007).

Project ScienceSpace is a VR training system that uses controlled experiments to provide a better understanding of scientific concepts. A virtual environment, *NewtonWorld*, allows students to experience Newton's three laws in an artificial environment that they can control. *MaxwellWorld* enables the examination of electrostatic forces and fields and *PaulingWorld* allows students to examine the structure of molecules (NATO, 2003).

The *Institute for Creative Technologies* at the University of Southern California developed a *Virtual Classroom VR System* (Rizzo, Klimchuk, Mitura, Bowerly, Buckwalter & Parsons, 2006). The original intention of the system was to provide a controlled stimulus environment to assess students who had attention-deficit or hyperactivity disorder. Following the success achieved in using the system, *Virtual Classroom* is also now applied to assess social anxiety disorder, eye movement under distraction conditions, and for earthquake safety training.

"A practical e-training development approach can facilitate and promote the development of competencies and knowledge in industry" (Sarraipa, Gomes-de-Oliveira, Marques-Lucena, Jardim-Goncalves & Mendonça da Silva, 2013:1). In a mining context, a primary aim of developing virtual environments is to give mine personnel opportunities to practise and experience 'authentic' situations, activities and processes, of the kind encountered in the day-to-day operations at a mining site. Safe and efficient planning and production are fundamental to profitable operations, and VR provides an intuitive means of exploring the diverse information associated with various processes (Van Wyk & De Villiers, 2009). VR training systems are discussed in Section 2.5.3.

2.4.3. Automotive industry

Driving simulators, similar to flight simulators, have become very popular. Several simulator configurations are available, ranging from a single computer screen with a games-type steering wheel to fully-immersive systems with a 360° field of view. Some systems use a replica of the actual dashboard or driving cabin of the vehicle as an interaction mechanism. The software can usually be configured for a variety of driving conditions, subjecting the trainee to driving scenarios covering varying levels of complexity. VR simulators are available for cars and trucks, and even for forklifts (5DT, 2012).

In November 2014, *Volvo* became the first company to offer virtual reality test drives of their new vehicles on *Google Cardboard*, a cardboard device that converts smartphones into functional VR headsets using two lenses and a magnet. The success of this campaign led *Volvo*'s senior vice-president to state that "Virtual reality is becoming the new way of creating content and storytelling" (Dua, 2014:1).

2.4.4. Aeronautics and aviation

Using highly sophisticated flight simulators, pilots practise responses to mechanical failures and other unusual situations. The reasons for using such simulators are increased safety and lower costs. Contemporary simulators are sufficiently sophisticated to allow complete training in complex manoeuvres, such as landing an aircraft, and this technology is deemed effective as flight training for many complex piloting tasks (Dawson, Meyer, Lee & Pevec, 2007). By incorporating satellite images, developers can increase the quality of aviation simulators. Simulators are also used for integrated crew training in normal and emergency operating procedures.

According to Lee (2005), many pilots indicate that the experience of flying current simulators approaches reality. This highlights that the effectiveness of a simulator depends on the perceived fidelity between the real aircraft and the flight simulator.

Bowling, Kaewkuekool, Khasawneh, Jiang and Gramopadhye (2003) developed a VR aircraft inspection system to improve the quality and reliability of aircraft inspection. The inspection regulations consist of several interrelated human and machine components. Moreover, visual inspection plays a significant role in ensuring aircraft safety. The results of such use of VR indicate that the inspectors' performance actually improved.

Another example of a VR training system in the aviation industry is the *Air Traffic Control Training Simulator*. The trainee experiences a virtual world that closely resembles a real air traffic control environment, including aircraft, radar, an airport and airport ground vehicles (Tata, 2014).

2.4.5. Military

It is clear from the literature that VR technology is of great interest to the military. During the past ten years, the number and variety of virtual reality applications in the military environment have increased greatly, involving various training applications, design of weapon systems, and mission preparation and execution (IITSEC, 2014).

The most important military application for VR is training. VR training can reduce cost and risk of casualties (NATO, 2003). Application areas include mission rehearsal, remotely-operated systems, practise of military medical procedures, tactical training, flight simulators and tank simulators.

Soldiers need to learn to handle potentially threatening situations that may require interactions with civilians. The United States has been developing realistic, intelligent avatars that act as civilians when VR is used to train soldiers for peacekeeping missions, manning of checkpoints and other contingency operations.

The following are specific examples of VR training in the military:

- The *European Aeronautics Defence and Space Company* has developed an advanced VR training simulator for the German Air Force. During the training, two trainees – the pilot and the gunner – wear head-mounted displays that provide stereoscopic views of the entire virtual battle scenario. The trainees have the task of identifying and engaging enemy objects in the virtual scenario. The facilitator or trainer can create new scenarios and monitor the simulation when in progress. All events during the simulation are recorded so that they can be evaluated subsequently (EADS, 2006).

- *Fifth Dimension Technologies* have developed a number of VR training tools for military training. The *Air Defense Training Simulator* is used to train operators of surface-to-air missiles. The *Rocket Launcher* uses immersive VR to teach the trainee accurate range estimates when launching rockets at stationary and moving targets. The *Polyphem Fiber Optic Guided Missile Training Simulator* allows the trainee to control a simulated fibre guided missile. The trainee learns how to control the missile and how to select the optimum targets during a virtual mission (5DT, 2012).

- The US Army recently conducted a training exercise, called *Squad Overmatch*, that supplemented current warrior skills training with situational awareness and resilience skills training using a combination of virtual and live environments. Due to the success of this integrated training, combined virtual and live training is expected to play a key role in future army training (IITSEC, 2014).

2.4.6. Electrical environments

The control of electrical systems with constantly increasing power needs, complexity, and automation, requires highly skilled maintenance. VR technologies are employed to train operators in the assembly, servicing and maintenance of high voltage equipment.

The use of realistic virtual learning programs within the practice of maintenance procedures, supports vocational training. Arendarski, Termath and Mecking (2008) investigated the use of VR for training in safety procedures related to an extra-high voltage power transformer. A virtual environment of the transformer is created with all the necessary details to assure the mechanical and electrical coherence of the model and the particular functionalities of each element in the system.

Trainees receive a particular mission to accomplish by performing certain tasks in a particular sequence. An example of a task in this virtual scenario is the unplugging of the transformer from the power net and the disconnection of important elements. Because of the high risks related to the procedure, it is very important to show the correct sequential order of the work processes, while paying attention to the safety rules so as to guarantee the security conditions.

In each training mission, four interaction modes are available.

- Discovery mode: The trainee can explore the virtual environment where the tasks must be accomplished.
- Presentation mode: The trainee is shown the correct procedures.

- Guided mode: The trainee must accomplish the tasks him/herself, but with assistance from the system.
- Free mode: The trainee must accomplish the tasks without any support.

Another example of virtual reality training in electrical environments is a high voltage line inspection training simulator, where a trainee inspects a virtual transmission line that includes both good and defective parts in order to identify and categorise defective parts.

2.4.7. Marine

The *Warsash Maritime Acad*emy at *Southampton Solent University* runs several simulators to enhance safety and improve efficiency.

- *Bridge Simulator:* This is used for practising passage planning and emergency responses. Three bridge simulators can be coupled to perform multi-ship manoeuvres.
- *Engine Room Simulator*: This is used for the development of systems management and team working skills.
- *Liquid Cargo-handling Simulator:* This is used for practising the handling of potentially dangerous bulk liquid cargoes such as oil, gas and chemicals.

These advanced simulators can also explore 'what if' scenarios in complete safety. The ocean simulation provides a variety of tides and currents, as well as varying visibility and fog conditions (Warsash, 2010). Another marine VR example is the *Unmanned Underwater Vehicle Training Simulator (UUV)*. In this simulator, trainees can practise operating and controlling a UUV, which is used in mine hunting. It also teaches the trainee the correct mine hunting procedures and mine identification skills (5DT, 2012).

2.4.8. Medical

Medical training institutes recognise that VR simulators can contribute considerably to improving the performance of medical professionals and the enhancement of patient safety (Ziv, Ben-David & Ziv, 2005). McGaghie (2007) highlights the fact that, apart from the usual benefits of VR simulation, the use of such simulators in medical training increases public trust in the profession.

A number of highly successful VR training applications have been developed since the 1990s, for example, systems that train physicians to palpate subsurface tumours (Dinsmore, Langrana, Burdea & Ladeji, 1997) and VR training for the diagnosis of prostate cancer (Burdea, Patounakis & Popescu, 1998).

VR is also used for learning routine medical procedures that can be difficult for novice practitioners to master, without causing discomfort and pain to patients. For example, anaesthesiology students can be trained to inject analgesic drugs into the epidural space by manipulating an instrumented needle on a virtual patient (Weiss & Jessel, 1998).

Doctors can also be trained in intricate surgical procedures by the use of VR. In minimally invasive surgeries, such as endoscopies and laparoscopies, VR can provide a view of the surgical fields that are normally blocked during such procedures in real life (Taffinder, Sutton, Fishwick, McManus & Darzi, 1998).

In contrast to the use of animals or cadavers which can only be dissected once, VR simulators provide medical students with a dynamic medium for the study of human anatomy and physiology. Using VR, students can access libraries of 3D images of healthy and pathologic body tissue, and can then view body parts from different perspectives while performing dissections (Fasel, Gingins, Kalra, Magnenat-Thalmann, Baur, Cuttat, Muster & Gailloud, 1997). At *Emory University Hospital,* a VR system was designed to train physicians to thread a catheter through a virtual circulatory system and view angiograms of the virtual patient. Such systems can be used to evaluate the ability of doctors to perform the tasks. The Director of Vascular Intervention at the hospital, Dr. Christopher Cates, commented on the system: "There is mounting evidence that virtual reality training is a better, faster and safer way for physicians to learn endovascular procedures than the traditional training" (Emory, 2004:2).

The following are further examples of VR training systems in the medical field.

- *Bronchoscope Training Simulator:* This enables a trainee to exercise bronchoscope navigational and procedural skills on a virtual patient. The system combines a physical transparent model of the lungs, where a real bronchoscope is inserted, with a computer graphics model which displays the images that would normally be seen with a real bronchoscope in a real patient.
- *Gastroscope Training Simulator:* This uses the same technology as the bronchoscope simulator, except that a gastroscope is inserted in a model of the stomach (5DT, 2012).

Due to the perceived value of VR training in the medical profession, Kneebone, Arora, King, Bello, Sevdalis, Kassab, Aggarwal, Darzi and Nestel (2010) propose the use of distributed simulations. These are low-cost, portable VR systems that can be accessed by healthcare professionals who do not have access to full-immersive simulation centres.

2.5. Applications of virtual reality in the mining industry

Mining in the 21st century is a high-technology industry. Mining companies strive to increase and maintain production, while simultaneously remaining competitive within the global economy. At the same time they also need to ensure their workers' safety and maintain a good safety record. The use of virtual reality offers mines an opportunity to develop tools and systems for a variety of purposes that can improve knowledge and understanding of the work environment. Virtual reality systems for the mining industry have been developed by many organisations with varying degrees of success (Stothard & Van der Hengel, 2009).

2.5.1. Mine planning and design

Mines need to improve safety, decrease operating costs, and increase product quality. VR technology can assist in improving the design of new equipment and processes. One of the products available to enhance mine design is *Abaqus*, which uses finite element analysis to simulate a mine's lifecycle in 3D. This enables engineers to evaluate safety and to improve design planning, implementation and operations (Oosthuizen, 2009).

The North-American based *Mining Innovation Rehabilitation and Applied Research Corporation* (MIRARCO) and *Laurentian University* jointly developed a virtual reality laboratory, an immersive facility that supports group work for data visualisation. As shown in Figure 2.6, the laboratory has a spherical screen of 3.9 million pixels to display seamless stereographic images. These three-dimensional images are projected onto the curved screen and up to twenty people can view these images in a theatre environment that enables team interpretation of the images. Mining companies bring their data to the laboratory to view a 3D virtual representation of their most complex datasets. These include data on mine geometry, geochemistry, geology and geomechanics. New exploration targets can be identified and decision-makers can understand the impacts of their decisions on mine planning, exploration, design and operations (Jamasmie, 2009; Kaiser, Henning, Cotesta & Dasys, 2002).

At the *University of Oviedo* in Spain, a virtual reality model is used to predict the response of longwall coal mining installations against changing operating conditions. This response is modelled using extensive data obtained from deep measurement campaigns and includes information on relevant geology, rock mechanics, stress-deformation calculations and hydraulics. Fuzzy logic, neural networks and 3D finite element analysis are combined with modelling tools to develop the model (Torano, Diego, Menendez & Gent, 2008).

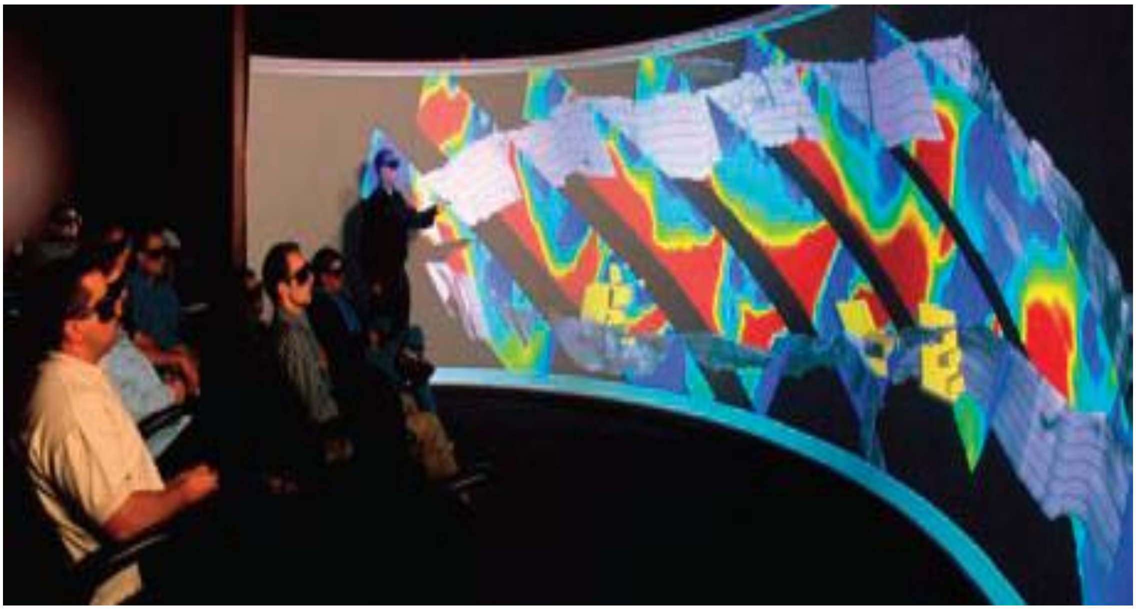


Figure 2.6: The Virtual Reality Laboratory at Laurentian University (Jamasmie, 2009:3).

In recent years a number of researchers in China have started investigating the use of VR in mining. One such example is the research of Mingming and Keping (2009), which led to the development of a virtual reality simulation for the underground mining processes of polymetallic ore deposits. The system supports mine planning and design by simulating the mine development area and transportation system in an interactive virtual environment.

2.5.2. Mining equipment

The application of 3D CAD (Computer-aided drafting) visualisation and virtual reality in the mining and mineral processing industry has been proven to enhance the interaction with clients, while saving time and money on overall projects (Delabbio, Hitchcock, Iturregui, Dunn & Eastick, 2002). The operational requirements of mining equipment are increasing due to increased demands regarding safety, quality and lower costs. Design of such equipment requires technical and operational expertise, based on the requirements of the particular mining process. 3D modelling of the equipment and the working environment is therefore a key factor in the design of mining equipment. In addition to 3D CAD, VR technology can be utilised to further enhance visualisation of the equipment in virtual environments that represent the actual working environments.

An example of equipment design in the mining industry is the design of underground vehicles. One of the vehicles used extensively is the *Loader Hauler Dumper* (LHD). The design of LHDs, in the context of constraints within the underground operating environment, has resulted in a number of serious accidents and fatalities, mainly due to the driver's restricted line of sight. It is believed that VR can add considerable benefit to line-of-sight reviews during the design process. Research work being undertaken by *MIRARCO* and *Laurentian University* addresses the need to improve the visibility of operators driving mobile equipment in underground environments. A VR model of LHDs in different mine layouts is helping to improve operator visibility in future designs (Delabbio, Dunn, Iturregui & Hitchcock, 2003).

Foster and Burton (2006) also investigated the use of VR regarding sightlines of underground vehicles, but with the intention of modelling potential sightline improvements to existing vehicle designs. The use of VR modelling enables operators and engineers to visualise and evaluate the potential benefits of retrofit improvements to vehicles before they commit themselves to the expense of modifying the vehicles. The retrofit improvements are modelled and the modified vehicles are operated within the virtual environment to observe the effects of modifications aimed at improving visibility. Examples of cost-effective retrofit improvements that have improved visibility by up to 31.4%, are chamfering of the engine cover and repositioning of the seat and headlight.

In another ergonomics study, Foster and Burton (2004) used VR technology to investigate ergonomic limitations associated with the use of the remote-controlled continuous miner. A continuous miner is a machine with a large rotating steel drum equipped with tungsten carbide teeth that produces a constant flow of ore from the working face of the mine. Remote-controlled continuous miners are used to work in a variety of difficult mining conditions. Pitzer (2000) found that several accidents using the continuous miner were related to a combination of poor operator positioning and inaccurate assessment or perception of risk.

In terms of ergonomics, Foster and Burton (2004) identified two categories of hazards:

- Hazards associated with the operator and his positioning, including being struck by moving machinery or by falling objects, as well as exposure to dust and noise.
- Hazards related to the use of the remote control, such as the operator activating controls incorrectly or in the wrong direction, resulting in him being injured by the continuous miner.

Based on this information, a VR model was developed to demonstrate the risks associated with the positioning of operators of continuous miners. A qualitative risk assessment was done and the information was used in the virtual environment to make operators aware of potential hazards and risks.

2.5.3. Training applications

Even though VR training is considered by many industries to be a safe and cost-effective alternative to face-to-face training by lecturers, the mining industry has been relatively slow to adopt such opportunities (Stothard & Van der Hengel, 2009). Acceptance is slow despite implementations of interactive training simulations such as those described by Unger and Mallet (2007), and despite the fact that VR was introduced to the mining industry in the late 1990s, when low cost simulators were discussed by Bise (1997) and Denby, Schofield, McClarmon, Williams and Walsha (1998).

Within the mining industry in South Africa, Squelch (2001) did pioneering work on developing a VR simulation system for a gold mine. He concluded that "the successful development of a VR underground simulator is possible and has demonstrated the potential for the application of VR training systems in the mining industry over the next decade" (Squelch, 2001:215).

To date, many of the developed systems have placed the emphasis on synthetic virtual models of the mine environment. Examples showing the development of such mining training simulations are presented by Denby and Schofield (1999); Filigenzi, Orr and Ruff (2000); Kizil and Joy (2001); Nasios (2001); Schofield, Denby and Hollands, (2001); Squelch (2001); Stothard, Otto, Laurence, Galvin and Zenari (2001); Orr, Filigenzi and Ruff (2002); Schafrik, Karmis and Agioutantis (2003); Schmid (2003); Kizil (2003); Schmid and Bracher (2004); Van Wyk (2006); Stothard (2007); Schmid and Winkler (2008); Van Wyk and De Villiers (2009); Ren, Kong and Ren (2012); and Yu and Chen (2014).

The next subsections overview different types of VR training systems developed for the mining industry.

2.5.3.1. Accident reconstruction simulations

Accidents in the mining industry are unfortunately a regular occurrence, due to the inherently hazardous nature of mining. With the aim of preventing recurrences, VR can be used to simulate the circumstances relating to previous serious accidents. Simulating a range of accident scenarios on a computer screen and viewing them from any angle enables accident investigators and workers to understand the underlying causes of an accident (Nasios, 2001).

Accident reports are usually highly technical in nature and not suitable as a means of mass communication. Even if a summary of accident findings is distributed throughout the company or industry, it can convey the details of the situation inadequately. Using VR techniques to reconstruct details of an accident, provides a powerful means of presentation. The simulation of events leading up to, and during, an accident can also be used as an integral part of the accident investigation. Viewing the scenario from different perspectives can serve to resolve conflicting reports from witnesses.

Accident reconstructions help to emphasise the significance of unsafe acts and promote a strong safety culture (Schafrik *et al.*, 2003). It enables the workers to understand how and why an accident happened, how it could have been prevented, and how injuries or fatalities could have been avoided (Kizil & Joy, 2001).

When an accident/incident occurs on a mine, it is followed by a series of investigations and, in the context of the South African mining industry, normally starts with what is termed a *Section 54*, being served on the mine. According to Section 54 of the Mine Health and Safety Act (MHSA, 1996), an inspector may halt the operations at a mine or part of a mine, should the inspector believe that any occurrence, practice or condition at that mine might endanger the health or safety of any person. In attempts to substantially reduce the injury rate, government inspectors are issuing *Section 54* notices to mines for any serious incident, or even potentially dangerous conditions noticed during inspection visits. This leads to serious production losses, as well as suffering on the part of the accident victims and their families. Incidents where losses occurred are thoroughly investigated and usually result in preventative actions being implemented. In most cases these actions are reactive, and it is frequently determined that the incidents could have been prevented by a rigorous risk management process. Such a process can identify major risks and management can take preventative measures to avoid recurrences (Webber-Youngman & Van Wyk, 2011).

Schafrik *et al.* (2003) noted that continuous employee education and training, as a means of establishing a strong safety culture in the industry, is a major factor in preventing fatal and non-fatal mining incidents. As Webber-Youngman and Van Wyk (2011) point out, VR training can play a meaningful role in the transfer of this knowledge.

Fowle (2003) focused on the development and reconstruction of incidents to ensure a greater understanding of the incidents and the events leading up to them. This led to the development of a training application with the view to improving safety standards and reducing accidents and incidents.

An animated *accident reconstruction simulation* comprises the following components (Van Wyk & De Villiers, 2009):

- A virtual environment with animated scenes depicting details regarding what exactly occurred.
- Scenes indicating the cause/s of the accident by highlighting the erroneous actions undertaken.
- Scenes indicating the correct procedures for such circumstances.

In South Africa, *Simulated Training Solutions (STS)* have simulated mining-related accidents (STS, 2013). In a recent instance, a speeding driver jumped from an out-of-control water carrier, after which the vehicle overturned and killed him. The exact circumstances of the accident were reconstructed by modelling the environment in 3D detail, using animation to indicate the causes of the accident and to demonstrate how the fatality occurred. The simulation can be run both in slow motion and as a bird's eye view so as to gain a better understanding of the sequence of events. The causes of the accident are then shown, as well as the correct procedures that should have been followed. In this particular example, the simulation was sent to the mine's head office in Switzerland to provide a further detailed explanation to senior management.

Figures 2.7A through 2.7D are screenshots from the reconstruction simulation that were provided to the researcher by STS. They illustrate the sequence of events. In Figure 2.7A, the speeding water carrier skipped the stop street. The driver lost control and the carrier headed towards the deep excavated area adjacent to the road (Figure 2.7B).

Probably fearing that the vehicle would fall into the excavated area below, the driver jumped out, but the vehicle hit a sidewall and overturned, as shown in Figure 2.7C.

Figure 2.7A: The speeding water carrier.

Figure 2.7B: The carrier heads towards a high wall.

Figure 2.7C: The water carrier hits a sidewall and starts to overturn.

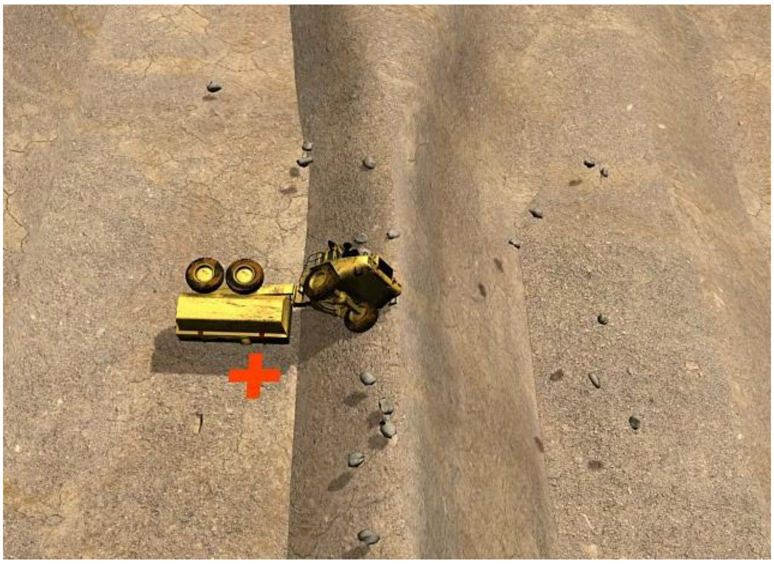


Figure 2.7D: A red cross indicates where the deceased was found after the accident.

Accident reconstruction simulations have also been used for training purposes as an attempt to prevent recurrences of such accidents.

The example depicted in Figures 2.8A through 2.8D, also supplied by STS, relates to a ground fall incident, where workers installed a safety net in the wrong area and then commenced work in the unprotected area. This resulted in a fall of ground and severe injuries to these workers. After examining and barring down loose rocks in the hanging wall, the workers identified intersecting joints (Figure 2.8A). A safety net was installed over a section of the hanging wall, but it did not cover the complete intersecting joint area (Figure 2.8B).

Figure 2.8A: Workers identified a number of intersecting joints.

Figure 2.8B: Workers installing the safety net.

The workers proceeded to install a roof bolt in the hanging wall, working in the unprotected area next to the safety net (Figure 2.8C). While they were realigning the drill, a section of the hanging wall fell and injured them (Figure 2.8D).

Figure 2.8C: Installation of a roof bolt in the hanging wall.

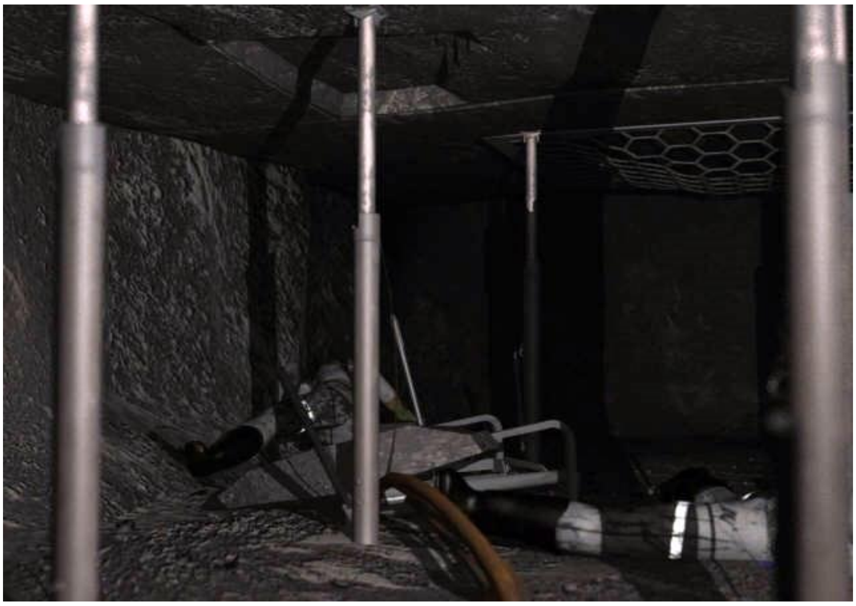


Figure 2.8D: Workers injured due to a rock fall.

2.5.3.2. Spotting generic hazards

VR applications for training in hazard awareness enhance traditional training methods without unnecessarily exposing trainees to actual hazardous situations. The *University of Nottingham* pioneered the use of VR training simulations for the mining industry with the development of a hazard identification simulation program (Denby and Schofield, 1999). The user could navigate around a virtual mine, carefully looking around to spot hazards, yet without any risk.

Kizil (2003) discusses the use of hazard walkthroughs as part of a quantitative risk assessment procedure, whereby a trainee can carry out a virtual inspection of the workplace, identifying hazards and assessing risk. An example is a virtual barring simulation system to provide improved hazard identification training for underground workers. The system focuses primarily on barring down exercises in rock-related hazards. The trainee is exposed to various hazardous situations and is required to navigate around the virtual environment, while identifying the hazards and selecting appropriate corrective actions.

Filigenzi *et al.* (2000) believe that the use of VR training tools could assist in reducing the injury and fatality numbers by allowing mine workers to practise identification and assessment skills from the safety of a computer. Researchers at the *National Institute for Occupational Safety and Health* (NIOSH) developed VR software to train mine workers on hazards identification, evacuation routes and evacuation procedures (Orr *et* *al.*, 2002). In evacuation training, the software enabled trainees to practise escape routes in a simulation of an emergency situation where fire, smoke and other dangers were present. By repeatedly practising the simulation, trainees become more familiar with emergency procedures and evacuation routes. First-person action games technology was used to produce a virtual mine environment, in which mine workers could be trained in hazard recognition and hazard avoidance.

Li and Kang (2014) explained the development of a VR-based system for coal mining safety training. This training system provides interactive exercises for learning self-rescue and escape after disasters.

Stothard and Laurence (2014) describe the use of VR simulations to visualise the impact of sustainable mining concepts of managing a mine site surrounded by a national park. The system incorporates historic and current data from the mine site, an environmental impact statement and environmental issues that must be resolved to the satisfaction of all stakeholders.

Van Wyk and De Villiers (2009) report on the use of VR for the development of an interactive *Plant Safety Program,* in which the present researcher participated. This training sensitises the employee to hazards in smelting plant areas. It is also applicable to visitors to plants, as it focuses on the dangers of areas that may be visited. The training includes simulations of incidents and hazardous situations at plants and assists trainees in the practical application of newly-attained conceptual knowledge of best practices and physical condition requirements in such a plant. Figure 2.9 shows an employee removing the protective conveyor drum guard without first stopping and locking out the conveyor, as is required by the safety procedure. During lockout, electrical systems are shut down to prevent the release of hazardous energy, and access to the switches is prevented by attaching a lock to the switch cabinet. The scenario in Figure 2.9 is an example of augmented reality, where computer animation is overlaid on video footage of the actual conveyor.

Figure 2.9: Worker removing the drum guard while conveyor is running (Van Wyk & De Villiers, 2009).

The *Plant Safety Program* uses a traffic light as metaphor. The objective is for the trainee to correctly identify why it is a 'red' condition and to change the situation to 'green' by correctly answering questions on how to deal with a particular scenario. A virtual environment portraying a specific substandard condition is shown to the trainee and he/she must correctly identify the condition by choosing the right answer from the available options. The traffic light can be turned to green by correctly identifying the substandard condition and by specifying how it should be treated. (Van Wyk & De Villiers, 2009).

Figure 2.10 is another example from the *Plant Safety Program* depicting a hazard relating to hot slag. Augmented reality is used to show the possible consequences of losing one's footing near the hot slag by overlaying computer-generated imagery of the worker falling into the hot slag over actual video footage of the slag.

Figure 2.10: Worker attempting to jump over hot slag (Van Wyk & De Villiers, 2009).

2.5.3.3. Training dealing with specific hazards

Various systems have been developed that focus on simulating hazards relating to specific machinery or particular situations in the workplace.

Lucas, Thabet and Worlikar (2007) discuss the use of VR simulations for training regarding hazards related to the use of conveyor belts. A training program was developed comprising a series of VR modules focusing on the components and assembly of conveyor belts and hazards associated with working or maintaining conveyors. The user needs to perform a number of pre-defined tasks in the virtual environment. The system offers two options for navigating the environment, in that users can select either an automated walkthrough or a manual walkthrough. In a subsequent article, Lucas and Thabet (2008), report on the use of VR for task-based training relating to conveyors. Learning can be enhanced by allowing the trainee to perform 'what if' scenarios. Using the VR system, consequences of actions can be experienced in a way that would not be allowed in the real environment. A task-based prototype was developed as a PC-based application, but it was also ported to an immersive CAVE environment and successfully used for training new recruits.

A conveyor hazards system was also developed in South Africa. *Simulated Training Solutions* (STS) produced a system to sensitise employees to hazards in and around conveyor systems (STS, 2013). Hazardous situations are simulated, as well as the potential consequences of ignoring them or working carelessly. A virtual workplace was created in which the system randomises hazards and the severity levels, which prevents monotonous repetitive exercises. The hazards include performing maintenance on conveyors without following the lock-out procedure, riding on conveyor belts, damaged belt structures, and cleaning procedures.

Another STS system focuses on the hazards associated with trackless moving machinery (TMM). This system highlights all the major hazards related to TMMs in an underground mechanised mining environment. A 'pedestrian' can be virtually placed in the simulated driver's seat, so that he/she can experience and understand the driver's limitations with regard to restricted field of vision, manoeuvrability and vehicle control (STS, 2013). Figures 2.11 and 2.12 are screenshots from a virtual hazardous situation, where a mine worker, contrary to safety regulations, requested an LHD driver to use the vehicle's bucket to lift him towards the hanging wall so he could replace a part on a fan. The driver accidentally lifted the bucket too high and crushed the mine worker against the hanging wall.

Figure 2.11: Mine worker inside the LHD bucket waiting to be lifted to the hanging wall (STS, 2013).

Figure 2.12: Mine worker crushed by the LHD bucket (STS, 2013).

Orr, Mallet and Margolis (2009) developed a VR system for enhanced fire escape training for mine workers. In the system, four trainees work together in a virtual mine via a computer network. Each trainee is represented by a computer-generated character and has independent control over his/her character in a first-person perspective on the virtual mine. The simulation represents a longwall coal mine in which, given various fire scenarios, trainees must locate an escape route. After successful trial runs, the evacuation training was included in compulsory annual safety training refresher courses.

Other examples of VR hazards training software for coal mines are discussed by Stothard *et al.* (2001) and Ji-zu, Li-mei, Jin-yun and Xiao-li (2009). The University of New South Wales developed a prototype where trainees had to spot hazards relating to an underground personnel carrier (Stothard *et al.*, 2001). VR technology can also be used for the simulation of hazards relating to the underground ventilation system (Ji-zu *et al.*, 2009).

In Australia, the *School of Mining Engineering* at the University of New South Wales and the *Australian Centre for Visual Technologies* based at the University of Adelaide are involved in a VR design and development partnership. Stothard and Van der Hengel (2009) describe a recent collaboration between the universities and a large mining company, *BHP Billiton,* regarding a VR safety training program for personnel at the *South Australian Olympic Dam* mine site. The program is called *Working at Heights* and is designed to familiarise the user with working at heights in different situations across a surface mine. Individual training modules of the program include operation of an elevated work platform, inspection of an open trench, inspection of scaffolding and the procedure for safely changing light globes using ladders. The program allows users to interact with safety documentation, equipment and procedures that they would encounter on site. The modules evolved through discussion with experienced industry trainers.

2.5.3.4. Training related to mining equipment

At the *Department of Energy and Geo-Environmental Engineering* at *Pennsylvania State University*, research was done on using VR for task training on mining equipment. In a study by Chakraborty and Bise (2000), workers were taught the basics of operating a continuous miner. The trainees were introduced in the virtual environment to the various controls and displays on the continuous miner, and could then familiarise themselves with the start-up procedure. The system also includes exercises on operating the continuous miner from the operator's compartment.

Many of the global mining equipment manufacturers have developed simulators of their more advanced equipment. *Sandvik* and *Atlas Copco* have drilling simulators and *Bucyrus International* introduced simulation training for electric mining shovels (Chadwick, 2009). The shovel simulation system is used as an introduction to safe and productive shovel operation and can also serve as refresher training for more seasoned operators. The trainee interacts with a simulated haul truck, while seated at the controls of a *Bucyrus* shovel in a virtual mine. These simulators allow training in the use of sophisticated equipment in the virtual environment, which is usually followed by a learning period on the actual equipment.

Caterpillar and *Volvo* supply simulators of their heavy machinery and equipment, including wheel loaders and excavators. Figure 2.13 is an example of a wheel loader simulator, where the operator sits in a cabin with the actual controls and operates the virtual machine at a virtual mine. These simulators feature state-of-the-art software with advanced 3D graphics to reproduce the operational movements of the real machines (CAT, 2014; Oryx, 2014).

Locally, two South African manufacturers offer VR training simulators to the mining industry. *Fifth Dimension Technologies* and *Thoroughtec* have simulators for roof bolters, continuous miners, LHDs and haul trucks, while *Thoroughtec* supplies mine trucks, rock drills and underground locomotive simulators (5DT, 2012; Thoroughtec, 2012).

Another example of a simulator is available from *Deutsche Steinkohle AG*, a German mining company that developed a simulator to perform underground activities such as cutting, bolting, loading and roof support. The VR software allows the user to navigate in all directions and to choose various camera positions to view the virtual environment (DSK, 2010).

Figure 2.13: Example of a wheel loader simulator (CAT, 2014).

Further examples of VR for mining equipment include the development of a training simulation of a remotely operated LHD vehicle to transport ore in hard rock mines (Swadling & Dudley, 2004), a VR training simulator tailored to the needs of the Australian coal mining industry where a continuous miner, a dump truck and a roof bolter were selected for a feasibility study (Stothard, Galvin & Fowler, 2004) and a VR system which provides a virtual training environment for operating drill jumbo's and loading machines (Schmid & Rossmann, 2004). *CAE Mining* has recently released its *CAE Terra* range of mining equipment simulators. These training simulators cover a variety of mining equipment types for both open pit and underground operations (International Mining, 2012).

2.5.3.5. Application to present study

Matthee, Henneke and Johnson (2014:42) highlighted the importance of training in the mining sector and commented that "in the face of current labour unrest and job cuts in this sector, it is foreseen that e-Learning might play an increasingly important role to upskill the remaining work force". Webber-Youngman (2014) indicated that virtual reality technology could be used to ensure the design of safe and productive mines.

With VR training being increasingly implemented in the mining sector, it is essential that prototypes and systems should be formatively and summatively evaluated. Despite many examples of VR training systems and simulators, the specific purpose, true capability and flexibility of such systems are not easily determined from the literature. Most publications in the commercial domain have only very broad descriptions with little reference to formal research literature in peer-reviewed journals that describe effectiveness, capability, findings of evaluations, and long-term benefits. Tichon and Burgess-Limerick (2011) examined the effectiveness of VR as a medium for safety-related training, including mining. Their report also cites the work of the present researcher (Van Wyk, 2006; Van Wyk & De Villiers, 2009), and concludes that "the use of virtual reality as medium for training in the mining sector is currently largely still at prototype stage, and rigorous and systematic evaluations have not been undertaken" (Tichon & Burgess-Limerick, 2011:27).

Zhang, Stothard and Kehoe (2010) and Bennett, Stothard and Kehoe (2010) have performed formal evaluations of the experiences gained by users of VR simulations in a mining context, each using their own set of criteria. The present study proposes a framework of criteria for the evaluation of desktop VR systems, specifically for the mining industry, which can be applied to evaluate the effectiveness of the instructional design, usability, VR system design and context-specific issues relating to mining.

2.6. Conclusion

This chapter presented virtual reality and its applications in various fields with specific emphasis on its use in training and the mining industry. Examples of different definitions of virtual reality found in literature were presented as an indication that the term *virtual reality* is used to describe various forms of visualisation technology, which can be confusing. Certain terminologies were defined, i.e. mixed reality, augmented reality and augmented virtuality. A taxonomy of interactive systems was presented to classify the different types of visualisation systems and the content that each class can display.

Three main categories of VR systems were defined in Section 2.2.1. This categorisation is based on levels of immersion: fully-immersive, semi-immersive, and desktop or non-immersive VR systems. This study explores the use of VR to improve safety training in mines, but it is important to note that the research will focus on using desktop VR systems which, at this stage, are the most appropriate tools to reach the high volumes of miners requiring safety training.

In Section 2.2.2, several important features of VR systems were introduced: computer-generated, three-dimensional, interaction, presence, viewpoints, realism and fidelity. These are key features relevant to the design of any VR system, and are included in the evaluation framework presented in Chapters Five and Nine. In the section on mixed reality, other terminologies were also defined, such as augmented reality and augmented virtuality. A taxonomy of interactive systems was presented to classify the different types of visualisation systems and the content that each class can display.

Section 2.3 provided many examples of applications of virtual reality technology in industry. Applications within the fields of engineering, science and medicine, entertainment and information visualisation were presented. The use of VR in training was highlighted as a separate section, Section 2.4, with examples from astronaut training, the use of VR in education, and VR training applied to the automotive, aeronautical and aviation industries, the military, as well as electrical, marine and medical environments.

The application of VR technology in the mining industry, which is the main focus of this study, was covered in the last part of this chapter, Section 2.5. VR is currently applied in the areas of mine design, mining equipment and training. The use of VR for training in the mining industry was discussed in the context of accident reconstruction simulations, general and specific hazardous environments and equipment simulators.

This chapter addressed Research Subquestion 1 of this study: "What is the suitability and potential of virtual reality technology for training applications in the domain of mine safety training?" in Sections 2.4 and 2.5. Virtual reality offers many possibilities in training and holds potential to increase productivity and improve safety awareness. Although the mining industry has been slow to invest in and use this advanced technology, the number of VR applications in the industry is increasing. Based on the amount of research currently undertaken on this topic, it is clear that the use of VR simulations in the mining industry will become more prevalent in the future. The hardware required to run non-immersive virtual reality systems is now available, even to home users, at an affordable price.

This chapter is the first of three literature studies. The next chapter, Chapter Three, discusses VR systems design, usability, and instructional design, while Chapter Four focuses on current safety practices in the South African mining industry.

Chapter Three

Systems Design, Usability and Instructional Design applied to E-learning environments

3.1. Introduction

This chapter discusses human computer interaction (HCI) aspects relevant to this research, namely systems design, usability, and instructional design. Section 3.2 covers different types of user-centred design methods: user-centred systems design, learner-centred design, interaction design and usability engineering. Section 3.3 develops the concept of usability, and includes subsections on usability of virtual environments and usability of e-learning.

As this study focuses on virtual reality training applications, the chapter also discusses relevant learning theories in Section 3.4. Section 3.5 takes an in-depth look at instructional design, with subsections covering the psychological theory underpinning design, design of multimedia learning, and methodologies that facilitate learning. The last section of this chapter, Section 3.6, discusses various usability evaluation methods and concludes with heuristic evaluation, one of the primary research methods of this study.

Figure 3.1. shows the layout of the chapter.

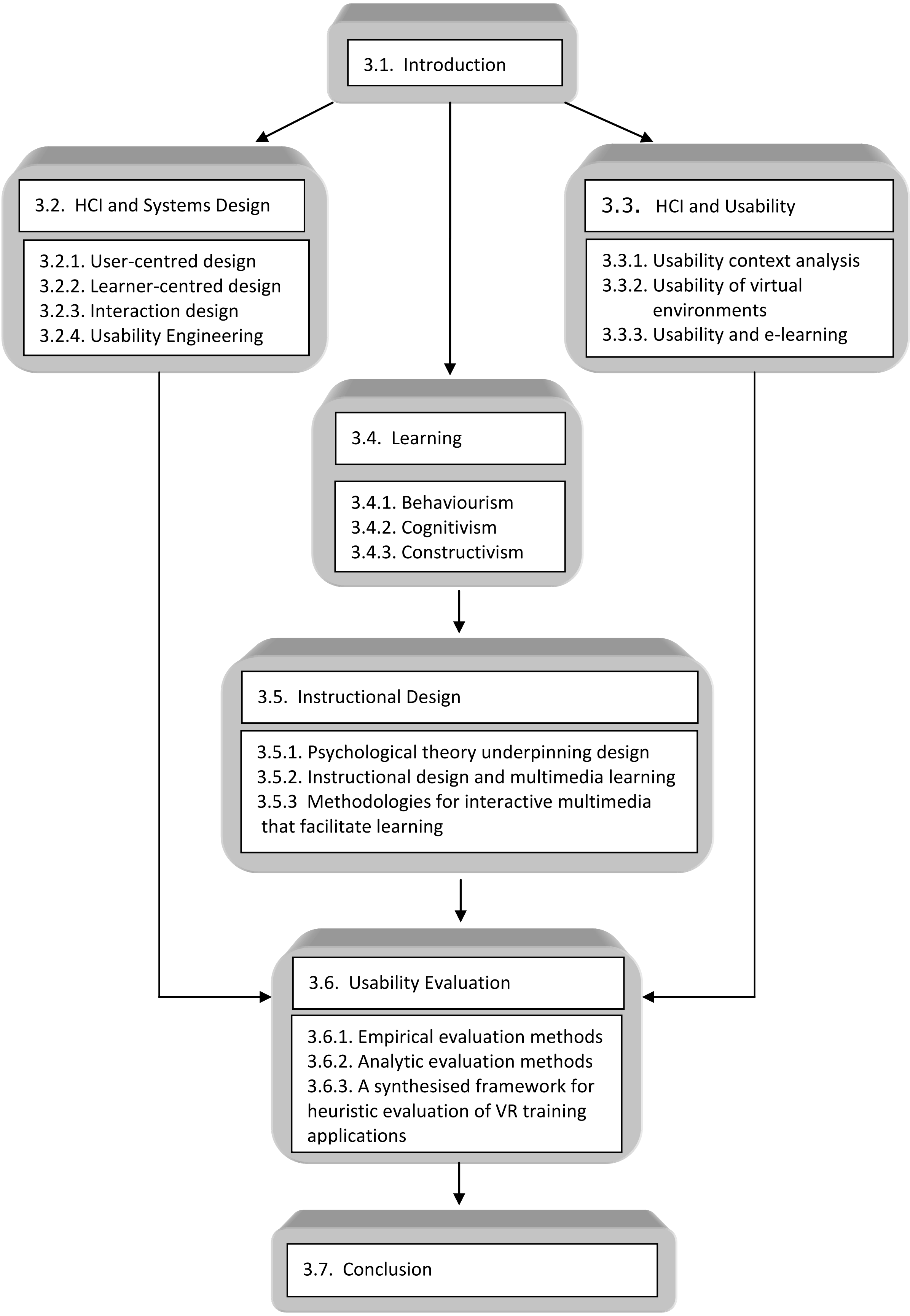


Figure 3.1: Layout of Chapter Three.

3.2. Human computer interaction and systems design

The Association for Computing Machinery (ACM) defines HCI as a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and the study of major phenomena surrounding them (ACM, 1996). The ACM defines four major areas within HCI, namely the use and context of computers, human characteristics, computer system and interface architecture, and the development process. To successfully develop an interactive system, knowledge is required regarding the intended users of the system, the behaviour and tasks of the users relating to their environments, technical possibilities, limitations and development tools, and processes or frameworks to guide design and development.

HCI provides a context in which to consider user-centred design methods and a basis on which to evaluate the efficiency and effectiveness of design methods. The following sections discuss different types of user-centred design methods.

3.2.1. User-centred systems design

According to Smith-Atakan (2006), traditional design methods tend to have a technology focus because user interfaces are often designed around a technical view of how systems work. In so doing, requirements, preferences, abilities and training needs of users can be overlooked. Further disadvantages apparent in traditional design methods, such as the Waterfall Model, include:

- The view of the designer is frequently reflected rather than the views of other important stakeholders. The user is often ignored or given a minimal role.
- Designers may find the system easy to use and may overlook critical design faults due to the 'familiarity paradox', which means the person who is most familiar with the system is often the least appropriate to evaluate it. When evaluating a system, designers may overlook an omission because they may hold some information in their minds which they automatically apply without realising it, or designers may find an item of information at a location because they know where to look, but it might be difficult for the user to find. A process that was developed and used many times by the designer may seem easy for the designer, but could prove to be difficult for the user.

In contrast, *User-Centred System Design (UCSD)* is a methodology that uses iterations of a three-phase cycle (Teixeira, Ferreira & Santos, 2012). Each cycle is informed by:

- requirements gathering through analysis of the users and their world,

- prototype design, and
- evaluation to ensure the design works well for users.

Norman and Draper (1986) were the first to use the term *User-Centred Design (UCD)*. They emphasised that the purpose of a system is to serve the user, and that the needs of the user should dominate the design of the interface. The importance of having a good understanding of the users was paramount, but they did not describe how to involve them actively in the design process. Karat (1997:33) defined UCD as an "adequate label under which to continue to gather our knowledge of how to develop usable systems". Emphasis was placed on involving users in system design without describing exactly how this should be accomplished.

Gulliksen, Goransson, Boivie, Blomkvist, Persson and Cajander (2003:401) defined UCSD as "a process focusing on usability throughout the entire development process and further throughout the system lifecycle". They developed a number of principles for the adoption of a user-centred development process, which they believe can be used to communicate the nature of UCSD, to develop processes that support a user-centred approach and to evaluate such development processes. Indicated below are some important principles.

- User focus: The users' goals, tasks, needs and context of use should guide the development.
- Active user involvement: Users should be directly involved and actively participate throughout the entire development process and system lifecycle.
- Evolutionary systems development: Feedback from continuous iterations with users leads to incremental software development.
- Simple design representations: Designs must be understood by all stakeholders.
- Continuous use of prototypes: When using prototypes of varying fidelity, users can visualise and evaluate design ideas.
- System use evaluated in context: Usability goals and design criteria should be used to evaluate designs.

Göransson (2004) also stressed that usability experts should be involved early and continuously throughout the development lifecycle.

Even though the basic principles and techniques are the same, there are different variations of user-centred system design processes (Henry, 2007). Figure 3.2 indicates the UCSD approach as advocated by Smith-Atakan (2006). The rectangles in the figure show activities, the black arrows indicate the sequence of activities and the blue arrows show the flow of information. Key to this approach is that users are involved in every stage. Figure 3.2 also indicates the outputs of each phase of the process. The output of each phase forms the input to the next phase in the process.

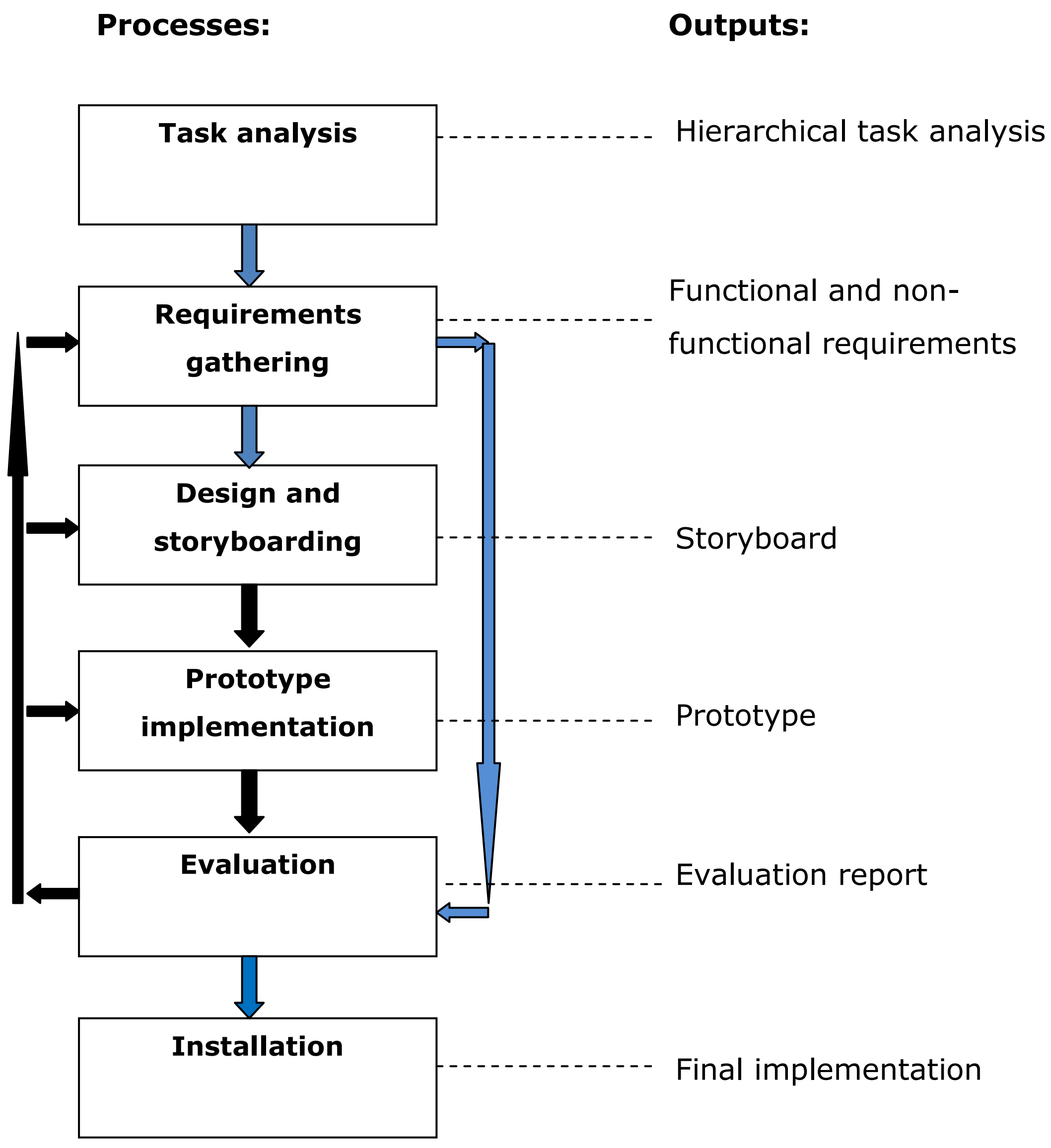


Figure 3.2: The User-centred system design process (Smith-Atakan, 2006).

UCSD processes are defined in ISO (International Organisation for Standardisation) documents, including ISO 13407 and the associated technical report, ISO TR 18529, proving that it is internationally endorsed as best practice (Mao, Vredenburg, Smith & Carey, 2005). User-centred design is also referred to as human-centred design. The following section briefly covers the key activities of UCSD, as depicted in Figure 3.2.

3.2.1.1. Task analysis

Task analysis provides a means of analysing and describing the tasks of users, so that they can be supported by interactive computer systems. It helps designers to understand existing systems and the users' existing tasks in order to better understand the user requirements of a system. The result of task analysis is a description of the tasks that users undertake when interacting with a system (Endsley & Jones, 2012). The inputs to task analysis are the problem statement and observations of existing systems and the output is an analysis in a hierarchical or matrix structure, called the hierarchical task analysis (HTA). An HTA involves identifying the goals that users want to achieve, decomposing these goals into tasks, further decomposing the tasks into subtasks, and this decomposition is repeated until the level of actions is reached (Smith-Atakan, 2006).

3.2.1.2. Requirements gathering

The purpose of this phase is to describe what the proposed system should do, without being concerned about how the system will support a task or how the system will appear. Inputs to this phase are the HTA, usability principles and factors such as technological limits or legal issues. The output is the requirements statement, consisting of functional requirements specifying the system functions and data requirements, and non-functional requirements which will include requirements relating to the environment, user groups and usability (Carroll, 2000; Teixeira *et al.*, 2012).

3.2.1.3. Design and storyboarding

This phase provides designers with an opportunity to visualise their designs and to review them in a fast and cost-effective way. Storyboards are created to indicate how the system will work and what it looks like. A storyboard is a hand-drawn mock-up of the system to be designed. Justification on why the system is going to work this way should also be included (Jantke & Knauf, 2005).

3.2.1.4. Prototype implementation

A prototype provides an early opportunity for users to evaluate a proposed system. The user interfaces look and behave like the complete system, but with limited functionality. Prototypes can be developed as throwaway (design ideas are carried forward into new developments and the prototype is discarded); evolutionary (the prototype is retained and more functionality is added); or incremental (the system is built as a set of separate components and each prototype is incrementally improved until it becomes a working component of the system). Prototypes can be of three types.

- Horizontal prototype: The user interface is simulated, but there is no functionality.
- Vertical prototype: There is full functionality for a limited vertical slice of the system.
- Full prototype: There is complete functionality, but with low performance (Smith-Atakan, 2006; Valk, 2007).

3.2.1.5. Design evaluation

UCSD is based on the belief that usable systems evolve through a process of generating, representing and testing ideas. These ideas are then refined during several iterations of this process. Such evaluation during the design process is also known as formative evaluation. Formative evaluation provides ways of learning about design options and points to strengths and weaknesses of proposed designs. This information forms the basis of new prototypes to be evaluated, and the process continues until a final design is accepted. Evaluation methods can include user observation, user experimentation, cognitive walkthroughs, heuristic evaluation, questionnaires and interviews, or combinations of these methods (Henry, 2007; Rogers, Sharp & Preece, 2011).

3.2.1.6. Installation

Installation is the final phase. By now the fully functional system with all its features has been through extensive evaluation and can be installed at the implementation site.

3.2.2. Learner-centred design

In 1996, Norman and Spohrer (1996:24) stated that “a revolution is taking place in education”, specifically referring to key terms such as constructivism, learner-centred, and problem-based learning, and indicating how the computer provides a powerful enabling technology for such philosophies. The new approach in the classroom, called *learner-centred,* was seen as similar to the *user-centred* focus of interface design, where the focus is on the interests, skills and needs of the learner. Three dimensions of instruction were identified.

- Engagement: Norman and Spohrer (1996) argue that a student who is engaged in the learning process is also a motivated student. Motivation is seen as the most important success factor. Interactive simulation technology, such as virtual reality, can contribute to motivating learners by providing compelling interaction and presenting information in forms that are easy to process perceptually.

- Effectiveness: Compelling interaction is of little value if learning does not take place. The system should be designed to ensure learning of the topics of concern.
- Viability: Social, cultural, political, technology and infrastructural issues can influence the viability of a system. Deploying a new pedagogy in an established training environment remains a major challenge.

Soloway and Pryor (1996) suggested that learner-centred design (LCD) represented a new generation of HCI that succeeded user-centred design (UCD) in educational contexts. The focus shifted from designing interfaces that support users in performing tasks to interfaces that support learning while performing tasks. They argue that students learn best when they engage in authentic, motivating tasks, and where scaffolding is applied. Scaffolding refers to a learning system that enables the learner to start doing a task with his or her current understanding, but that also channels, supports and challenges learners to develop the next level of understanding and performance.

Within the context of e-learning, learner-centred design focuses on the experiences, perspectives, backgrounds, capacities and needs of learners, as well as on teaching practices most effective in promoting motivation, learning and achievement (McCombs & Whisler, 1997). McCombs and Vakili (2005:1582) stated that "many researchers and practitioners are decrying the lack of a research-validated framework to guide their design" and proposed a learner-centred framework for the design of e-learning. The framework consists of 14 principles, categorised into four domains of learner-centred factors, shown in Table 3.1: cognitive and meta-cognitive factors, motivational and affective factors, developmental and social factors, and individual-differences factors. E-learning systems can also be evaluated by relating program features to the factor domains and specific principles.

Table 3.1: Learner-centred principles for designing e-learning (McCombs & Vakili, 2005:1586).

Domain	Principle
Cognitive and meta-cognitive factors	1. Nature of the learning process
	2. Goals of the learning process
	3. Construction of knowledge
	4. Strategic thinking
	5. Thinking about thinking
	6. Context of learning
Motivational and affective factors	7. Motivational and emotional influences on learning
	8. Intrinsic motivation to learn
	9. Effects of motivation on effort
Developmental and social factors	10. Developmental influences on learning
	11. Social influences on learning
Individual-differences factors	12. Individual differences in learning
	13. Learning and diversity
	14. Standards and assessment

3.2.3. Interaction design

Rogers *et al.* (2011) advocate the use of the term *interaction design* as an umbrella term covering interface design, user-centred design, web design, software design and interactive system design. Interaction design is not prescriptive in terms of methods, but promotes the use of a range of methods, techniques and frameworks. The main focus is on the design of user experiences. The process of interaction design involves four main activities:

- The needs of the user should be determined and the requirements for the user experience established.
- Alternative designs should be developed, based on the requirements.
- Interactive versions of the design should be developed.
- The resulting user experience should be evaluated.

Evaluation ensures that the end product is indeed usable. A user-centred approach to design is followed to involve users throughout the design process. In order to understand how to design interactive systems, Rogers *et al.* (2011) emphasise the fact that designers should understand the context in which the users live, work and learn. A number of general design principles are proposed to assist designers when designing user experiences.

- Visibility: The more visible functions are, the more likely it is that users will know what to do next.
- Feedback: Users should receive information on what action has been done and what has been accomplished.
- Constraints: The kinds of user interaction that can take place at a given moment should be restricted.
- Consistency: Interfaces should be designed to have similar operations and use similar elements for achieving similar tasks.
- Affordance: Objects and systems should be designed with attributes that support users in using them intuitively, for example, scroll bars should afford moving up and down and a mouse button affords pushing.

The emphasis in interaction design is on designing for the user experience and not just for usable products. In order to optimise the interaction between users and interactive products, designers need to consider a number of interdependent factors, including context of use, types of activity, cultural differences, user groups and usability goals.

Weiss (2008) argues, for example, that in the case of the design of a business website, interaction design should be extended to include marketing principles. Although interaction design is defined as multi-disciplinary, marketing expertise is not necessarily vital to interaction design. As an example, if a university website is designed using interaction design, an evaluation in terms of user experience may have a positive outcome in that the user finds the site easy to learn, enjoyable and useful. However, that does not mean that the university's goal of increasing its student numbers will be achieved. It is therefore proposed by Weiss that interaction design principles and marketing principles should be combined in what is termed *results-based interaction design*. In this way, the World Wide Web can be leveraged to its full potential in order to meet organisational goals, as well as provide a satisfactory experience for the user.

A similar approach to expand interaction design can be followed relating to interactive simulation training systems, as in this study. Trainees can be engaged in what they perceive as a motivating, useful, enjoyable and easy-to-use learning experience, but that does not guarantee that learning actually took place or that trainees will apply the principles that were taught. The design of the training system should also take pedagogical principles and learning theory into account.

3.2.4. Usability Engineering

Although the concept of usability will be introduced in the following section, another term used in literature to describe the process by which usability is ensured in interactive applications is *usability engineering*. Hix and Gabbard (2002) define the phases in usability engineering as user task analysis, user class analysis, design of the user interaction, rapid prototyping, user-centred evaluation and iterative re-design based on the evaluation results. Similar to the concepts addressed in the previous subsections, usability engineering involves users in the design and evaluation of a system (Prantosh, Kalyan, Dipak & Rajesh, 2014).

3.3. Human computer interaction and usability

The foundations of HCI focus on the psychological and physiological attributes of the human user, the capabilities and limitations of computing devices, and the dialogue between the two. HCI design practice addresses usability from the human perspective.

Usability is a general quality of the *appropriateness to a purpose* of an artefact (Brooke, 1998). This means the usability of any tool or system has to be viewed in terms of the context in which it is used, and its appropriateness to that context. In the ISO 9241 standard, usability is defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO, 1998). The context of use includes the users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used. According to ISO 9241, three potential ways of measuring the usability of a software product are:

- analysing the process of interaction, by modelling the interaction between a user and the system,
- analysing the effectiveness and efficiency, by measuring satisfaction of users of a product in a particular context, and
- analysing the features of the product, by assessing the features required of a product for a particular context of use (Avouris, 2001).

Usability ensures that interactive products are easy to learn, effective to use and enjoyable from the user's perspective (Jooste, Van Biljon & Mentz, 2014). Poppe, Rienks and Van Dijk (2007) state that the focus of HCI research has evolved over time in that the task- and work-related usability paradigm has been expanded to a holistic user experience.

Usability is thus regarded as ensuring that interactive systems are effective, enjoyable and easy to learn. Rogers *et al.* (2011) subdivide usability into various goals.

- Effectiveness: How good a system is at doing what it is intended to do.
- Efficiency: The way a system supports the users in performing their tasks.
- Safety: Protecting the users from dangerous conditions. (This is particularly relevant in hazardous conditions of work, such as mining).
- Utility: The extent to which the system provides the users with the functionality they require.
- Learnability: How easy a system is to learn to use.
- Memorability: How easily users can remember how to use the system, once learned.

The purpose of these usability goals is to provide designers with a means of assessing the user experience. By asking questions related to the goals, designers can be alerted early in the design process to potential design problems. Usability criteria can be derived from these goals, which can be used to assess the usability of a system. Such usability criteria provide quantitative information by measuring the extent of improvement, for example, the time a user takes to learn a system task (learnability) or the time taken to complete a task (efficiency). User experience goals can be set to obtain qualitative information on the user experience, for example, whether using a product or system is enjoyable, aesthetically pleasing, challenging, engaging, satisfying and motivating.

The International Organisation for Standardisation's ISO 9241 deals with the ergonomics of human computer interaction. Part 210 of this standard focuses on human-centred design for interactive systems and provides guidance on human-system interaction throughout the lifecycle of interactive systems. Regarding user experience, the standard defines it as "a person's perceptions and responses that result from the use or anticipated use of a product, system or service" (ISO, 2008). The three factors that influence user experience are listed as the system, the user and the context of use.

In the Software Quality standard ISO 9126, important attributes of software quality are functionality, reliability, efficiency, maintainability, portability and usability. Usability is further described as relating to other particular attributes.

- Understandability: To what extent does the user understand how to use the software and comprehend its conditions of use and suitability.

- Learnability: To what extent is the user supported in learning how to apply the product.
- Attractiveness: The attributes of the software that are attractive or engaging to the user.
- Operability: Capabilities of the software that enable user control and operation. This concept includes adaptability, changeability, installability and conformity to user expectations (Bevan, 1997).

ISO 9126 was extended in 2004 to include a section on *quality in use.* In 2011, ISO 9126 was replaced by ISO 25010, where usability is defined as both an *intrinsic product quality characteristic* and a *subset of quality-in-use*. As a product characteristic, usability has the intrinsic subcharacteristics of learnability, appropriateness, recognisability, user error protection, user interface aesthetics and accessibility. Quality-in-use refers to effectiveness, efficiency and satisfaction (Cockton, 2013).

Sachs (1995), Beyer and Holtzblatt (1998), and Harris and Henderson (1999) emphasise that computer systems in a workplace environment should support the particular work operational practices within the context of each situation, requiring a deep understanding of the context of use of the system. Gulliksen *et al.* (2003) believe that few development teams have sufficient knowledge of the contexts for which they design.

Merely compiling requirement specifications or creating abstract models is inadequate to create a sufficient understanding. This can only be provided by the users themselves. This argument is particularly relevant for the development of interactive training systems for the mining industry, where a sound sufficient knowledge is required regarding the intricacies of the context of use.

3.3.1. Usability context analysis

Usability context analysis (UCA) is a structured method for eliciting detailed information about a product and how it will be used, and for deriving a plan for a user-based evaluation of a product. UCA is discussed in detail in Chapter Six.

3.3.2. Usability of Virtual Environments

General usability and human factor issues of VR have been examined by several authors, including Galimberti and Belloni (2003), Sutcliffe and Kaur (2000), Wilson (1997) and

Bowman, Gabbard and Hix (2002). Further investigation, however, is required regarding aspects of participating in virtual environments (VEs) and their consequences for usability specific to training environments.

Until recently the focus of developers of virtual environments was largely on utilising the ability of the technology to provide a 'WOW' factor, and little attention was paid to usability issues. Hix and Gabbard (2002:681) remarked that "few principles for design of VE user interfaces exist, and almost none are empirically derived or validated". They present three usability engineering methods for VEs: user task analysis, heuristic evaluation, and formative usability engineering. These methods have been successfully applied in the development of other graphical user interfaces (GUIs) and have been adapted to be applied to VE development. They argue that a combination of these methods, as a progression from the three methods and followed by a summative evaluation, is an efficient and cost-effective usability engineering strategy for VEs. Each method generates information used for the next method, and by using more than one method more complete coverage of the usability issues is achieved. Use of more than one research method provides methodological triangulation (Cohen, Manion & Morrison, 2011).

3.3.3. Usability and e-learning

Some narrow definitions for e-learning define e-learning exclusively as using the Internet for instruction and learning, but other definitions are wider. Clark and Mayer (2003) define e-learning as "instruction delivered on a computer by way of CD-ROM, Internet or Intranet". Cedefop (2002) defines e-learning as "learning that is supported by information and communication technologies (ICT). e-Learning … may encompass multiple formats and hybrid technologies, in particular, the use of software, Internet, CD-ROM, online learning or any other electronic or interactive media". Mayes and De Freitas (2005:5) describe e-learning as "technology enhanced learning", which is the "use of technology to support and enhance learning practice". Sangrà, Vlachopoulos and Cabrera (2012) defines e-learning as an approach to teaching and learning that is based on the use of electronic media and devices as tools for improving access to training, communication and interaction, and that facilitates the adoption of new ways of understanding and developing learning.According to these definitions, the VR training solutions discussed in this study can be viewed as e-learning artefacts.

Mayes & Fowler (1999), writing more than a decade ago, highlighted how technological developments are changing perceptions of the learning task, making the need for effective design and evaluation approaches even greater. They argue that the usability of educational software cannot be measured in the same way as usability evaluation of conventional software designed for the workplace. This is because learning is a by-product of understanding, rather than an activity which can be supported directly. An e-learning application should therefore be pedagogically suitable, which means that both the tools, as well as the kind of interaction provided, must be aimed at supporting the learner in the specific learning task, rather than being a mere exercise of advanced technology.

Masemola and De Villiers (2006), Adebesin, Kotze and Gelderblom (2010), and Nyang'or, De Villiers and Ssemugabi (2013) point out that usability evaluation of e-learning differs from the evaluation of other software in a number of ways:

- Efficiency in e-learning cannot necessarily be judged by users being able to complete tasks quickly, as users have different learning styles and different approaches to working through the learning material.
- It is not always desirable to minimise errors in e-learning applications. Usability-related errors should be avoided, but cognitive errors could be part of the learning process and should be permitted where support is in place to help users recover from the error.
- In e-learning, the functional operations undertaken by users are learning activities, so the learning process is part of the instructional functionality.

This leads to the conclusion that "the effectiveness of learning and the users' subjective satisfaction with a resource are therefore part of its usability" (Masemola & De Villiers, 2006:188).

Squires (1999) indicates the importance of contextualising usability issues in terms of the complex tasks involved in learning. Squires and Preece (1999) investigated how usability features can be integrated with educational design to enable educators to evaluate educational software.

Usability plays an important role in the success of e-learning applications. If an e-learning system is not usable, learners will spend too much time on understanding the software functionality, rather than understanding the learning content (Costabile, De Marsico, Lanzilotti, Plantamura & Roselli, 2005). Ardito, De Marsico, Lanzilotti, Levialdi, Rossano and Tersigni (2004a) identify the adoption of a learner-centred design methodology as the key to developing systems conforming to usability criteria. UCD assumes similar experiences and common culture among users, but LCD considers a variety of different learning strategies, motivations and experiences.

An e-learning system should be pedagogically suitable, engaging and attractive (Ardito, Costabile, De Marsico, Lanzilotti, Levialdi, Plantamura, Roselli, Rossano & Tersigni, 2004b). Several authors stress the importance of context in the design and evaluation of e-learning (Jones, Scanlon, Tosunoglu, Morris, Ross, Butcher & Greenberg, 1999; Mayes & Fowler, 1999; Squires & Preece, 1999).

3.4. Learning theories

E-learning entails supporting learners in the process of learning. It involves information transfer rather than information translation, supporting human cognition, implementing behavioural change, and leveraging technology as a medium and messenger, rather than being a message or showpiece in its own right. "Foundations for e-learning must be based on sound principles of learning theory and instructional design, in order to facilitate effective learning" (De Villiers, 2005a:351).

According to Govindasamy (2002), one of the most important aspects of e-learning that is often neglected, is the need for careful consideration of the underlying pedagogy. Alessi and Trollip (2001) also emphasise the importance of assessing whether the design of an educational application reflects an underlying learning theory. It is therefore important to address the current learning theories on which e-learning are based.

Mayes and De Freitas (2005) suggest that there are no models specifically for e-learning, but there are e-enhancements of models of learning. This involves the application of technology to achieve better learning outcomes, to bring the learning environment to the learners in a more cost-efficient way, and to provide effective ways to assess the learning outcomes. Learning theories provide empirically-based information regarding the variables that influence the learning process.

Although no universal agreement exists on how learning takes place, psychologists and educators have generated several different principles and theories of learning. The following three are the main theories of the past three decades.

- Behavioural psychology: Learning is viewed as changes in the observable behaviour of the learner due to events in the environment,
- Cognitive psychology: A complete explanation of human learning also requires consideration of non-observable cognitive constructs, such as memory, mental processing and motivation.

- Constructivist approach: Each individual constructs his/her own view of reality and learners are active creators of knowledge, who learn by observing, manipulating, and interpreting the world around them (Alessi & Trollip, 2001).

3.4.1. Behaviourism

According to the behaviourist theory of learning, human behaviour is a product of *stimulus-response* interaction. Proponents of behaviourism maintain that the psychology of learning should be based on the study of observable behaviours and environmental events, and not non-observable constructs, such as memory or beliefs (Alessi & Trollip, 2001). This means that all complex behaviours, for example, emotional reactions and reasoning, are composed of simple stimulus-response events that can be seen and measured (Black, 1995).

Behavioural psychologists use incentives such as grades and tangible rewards to motivate learners to accomplish the educational requirements. With regard to teaching and learning, behaviourists suggest that the subject content of the curriculum should be sequentially organised, and learners assessed according to the standard they achieve (Black, 1995).

According to Alessi and Trollip (2001), one of the instructional technologies based on behaviourism is computer-assisted instruction (CAI). Traditional CAI typically comes in the form of drill-and-practice activities, simulations and tutorials. Behavioural principles, followed in sequence, can be used to program educational applications:

- Clearly express the purpose of the application.
- Thereafter, consider the most suitable multimedia, whether in visual or audio form or within the text, for the presentation of the content.
- Reward or 'punish'. After each question or exercise, 'rewards' reinforce or encourage positive responses, while an attempt is also made to minimise negative responses.
- Scoring monitors progress.
- Ensure that the status of the learner's progress is provided.

Those critical of the behaviourist theory maintain that unobservable aspects of learning, such as memory and motivation, thinking and reflection, are ignored. Furthermore, they maintain that too much emphasis is placed on the instructor and the instructional material rather than on the learner. Despite this criticism, behaviourism has had a substantial influence on teaching and learning and for many years provided the foundation from which many CAI applications were designed (Alessi & Trollip, 2001).

3.4.2. Cognitivism

Cognitive psychology is a branch of psychology focusing on mental processes, including how people perceive, think, remember and learn. In the cognitivist learning theory, emphasis is placed on unobservable mental constructs, such as memory, attitude, motivation, reflection, and other internal processes.

Learning, thinking, language, perception and reasoning are seen as outputs of an individual's attention, memory and concept formation processes (Mayes & De Freitas, 2005). There are two schools of thought in cognitivism:

- The *human information-processing* approach suggests that people use their senses for gaining information which is then stored in memory before being retained or forgotten. Initially, the information is stored in short-term memory, also known as the working memory. In order to be retained and stored in long-term memory, the information needs to be used or organised. The assumption is that the senses and the brain follow systematic, albeit complex, laws, and that learning can be facilitated in line with them.
- The *semantic network theory* maintains that the brain consists of billions of cells, or nodes, and that there are billions of links between them. These relationships are characterised by such aspects as similarity or opposition, cause and effect, or time. This theory claims that cognitive activities such as problem solving and acting, thinking and remembering, are activated by other nodes. Prior knowledge, according to this theory, is vital, and learning comes about when new knowledge is incorporated into the network of prior knowledge (Alessi & Trollip, 2001).

Included in the concept of the cognitive domain in cognitive psychology, is the recognition and recall of knowledge, as well as the development of understanding and intellectual skills and abilities. The ability of an individual to evaluate, reflect on, and manage his/her own cognitive skills is called metacognition. According to Reigeluth & Moore (1999), metacognition is an intellectual skill considered to be part of the cognitive domain. A high level of metacognition within an individual positively influences the learner's ability to learn, and vice versa.

Cognitive education seeks to improve methods of teaching and provides sets of instructional methods that assist learners in acquiring knowledge to be recalled or recognised, as well as developing learners' comprehension and intellectual abilities and skills. Reigeluth & Moore (1999) proposed four categories of how learning occurs, encompassing the concepts mentioned by other theorists on cognitive education, including Ausubel, Gagne, Anderson and Merrill. These categories are:

- memorisation of information,
- understanding relationships,
- application of skills, and
- application of generic skills.

3.4.3. Constructivism

The major theme behind the constructivist theory is that learning is an *active process* during which learners construct new ideas or concepts based upon their current and past knowledge (Bruner, 1990). Soloway, Jackson, Klem, Qumtan, Reed, Spitulnik, Stratford, Studer, Eng and Scala (1996:190) describe the constructivist view of learning and understanding as being "active, constructive, generative processes such as assimilation, augmentation, and self-reorganisation". As learning theories developed and instructional designers gained experience in computer-based technology, a shift of emphasis occurred from the behaviourist paradigm to the constructivist paradigm (Squires & Preece, 1999).

In the constructivist approach to learning, the emphasis is on learning being a distinctive and personal process, and that it is characterised by individuals who form and refine concepts and, in so doing, develop and interpret knowledge and understanding. In the learning environment, there should be many and varied knowledge representations and media, as well as cases and contexts. As learners explore systems, environments and artefacts, they learn to take responsibility as a result of the sense of ownership over their learning (Reeves & Reeves, 1997:60; Squires, 1999:464; Zhao & Deek, 2006:1589).

Proponents of constructivism point out that education has treated learners as passive vessels into which knowledge is poured. They propose that educators should rather take on the roles of coaches or facilitators of learners. Thus, designers of educational technology should aim to create environments that will facilitate the construction of knowledge. Various suggestions and principles have been put forward to assist in achieving that goal (Alessi & Trollip, 2001):

- The emphasis should be on learning rather than on teaching.
- The thoughts and actions of learners should be emphasised over those of

educators.

- Active learning is critical.
- Discovery or guided-discovery should be facilitated.
- Learners should be encouraged in constructing of information.
- Learning activities that require collaboration or cooperation have value.
- Focus on purposeful or authentic learning activities that are relevant to the learner.
- Personal autonomy on the part of learners should be encouraged.
- Learners need to reflect.
- Encourage learners to take ownership of their learning and activities.
- Learners should have opportunity to reflect on the complexity of the real world.

In line with the above principles, Jonassen (1994) describes learning environments that facilitate purposeful knowledge construction, as:

- Providing multiple representations of reality;
- Avoiding oversimplification of instruction by representing the natural complexity of the real world;
- Focusing on knowledge construction;
- Presenting authentic tasks by use of contextualised rather than abstract instruction;
- Providing real-world case-based learning environments, rather than predetermined instructional sequences;
- Enabling context- and content-dependent knowledge construction; and
- Supporting collaborative construction of knowledge through social negotiation, but not through competition between learners.

Squires and Preece (1999) stress the fact that learning should be authentic. Authentic learning can be considered from both a cognitive and a contextual perspective.

3.4.3.1. Cognitive authenticity

Learning that is cognitively authentic involves experiences where learners are assisted in constructing and refining concepts in personally meaningful ways. Squires and Preece (1999) identify three concepts resulting from cognitive authenticity.

- Credibility: Learners will experience credibility if they can explore the behaviour of environments or systems; if they receive intrinsic feedback from the environment; and if the environment provides a mechanism for learners to articulate ideas or opinions.

- Complexity: Learning environments should contain interesting and motivating tasks, which may lead to complex environments. Learners can be helped to cope with complexity by the provision of scaffolding, anchoring and problem-based environments.
- Ownership: Learners should be encouraged to take responsibility for learning. Strategies to encourage metacognition can lead to a sense of ownership.

3.4.3.2. Contextual authenticity

Cognition and learning are situated in specific learning contexts and all the components of a learning environment contribute to the learning process. This may require educators to guide learners to appropriate contexts, especially when learners need to move beyond understanding a concept in a specific context only, so that it can be applied more generally. The curriculum is also an important aspect of the learning context (Rogers *et al*., 2011).

3.4.4. Discussion of learning theories

E-learning applications reflect different views on cognition and learning, including behaviourism, cognitivism and constructivism. These learning perspectives provide structured foundations for planning and conducting instructional design activities.

Behaviourism's focus is on the external observation of lawful relations between and among outwardly observable stimuli and the responses that follow (Boghossian, 2006). In the traditional behaviourist model, learners undergo some form of conditioning. Ultimately, the goal of conditioning is to produce a behavioural consequence. As such, the primary responsibility of the instructional designer is to identify and sequence the contingencies that will help learners learn.

While the behavioural perspective has an external focus, the cognitivist approach has an internal one. During the 1980s, within the field of cognitive psychology, it became fashionable to discredit behavioural theories in learning. In 1994, Reeves (1994:225) indicated that "most 'self-respecting' instructional design theorists now claim to be cognitivists". Psychologists and educators began to place more emphasis on the role of cognitive processes in learning, such as thinking, problem solving, language, concept formation and information processing.

During the 1990s, however, some contemporary cognitive theorists began to adopt a more constructivist approach to learning and understanding. Ertmer and Newby (2013:55) report that "in recent years, constructivism has begun to receive increased attention in a number of different disciplines, including instructional design". There are many different types of constructivism, among the most popular are cognitive, critical, radical, and social. However, they all share the same core: the idea that learners construct their own knowledge (Sener, 1997). Constructing knowledge means that students are active participants in a learning process by seeking to find meaning in their experiences. In a literal sense, learners construct or find meaning in their subjective experiences, and this result becomes knowledge (Boghossian, 2006). Learning is therefore described as a change in knowledge stored in memory. As a consequence, the instructional designer is challenged with organising new information for presentation, carefully linking new information to previous knowledge and using a variety of techniques to guide and support the mental processes of the student.

Constructivism relates to personal knowledge construction and interpretation. It aims to "instil personal goals and active involvement within real-world situated learning, leading to application skills and transfer" (De Villiers, 2005a:359). Learner-centred environments are created, within which learners can explore. In summary, according to the constructivist approach to learning, learners should be given ownership of their learning, encouraged to explore, provided with meaningful real-world learning tasks, and should collaborate with educators and peers in order to discover and make meaning of new knowledge.

These learning theories are not necessarily mutually exclusive. The underlying theory behind any e-learning artefact can be a hybrid between two paradigms. For example, an instructional designer may define clearly an expected behaviour from a learner (behaviourist perspective) while he or she can establish a group activity or problem-based activity (constructivist perspective) by means of which the learner will practice the knowledge acquired.

3.5. Instructional design

Unlike conventional business applications where computer technology is applied to process business transactions and to generate concrete products, e-learning applications should support learners in the process of learning. According to De Villiers (2005a), this process involves information transfer, management of educational interaction, the support of human cognition, implementation of behavioural change, and leveraging technology as a medium or messenger and not as a message in its own right.

Ruffini (2000) describes *Instructional design* (ID) as the systematic planning and development of instruction. It involves a set of decision-making procedures by means of which the most effective instructional strategies are developed or chosen, given the outcomes learners are to achieve and the conditions under which they are to achieve them (Winn, 1990). The design of any instruction usually involves the use of instructional theories, design models and strategies, to help learners develop knowledge and skills (Dijkstra, 2001). Reigeluth (2013) describes instructional design as involving methods of instruction and contextualisation to the situations in which those methods should be used.

3.5.1. Psychological theory underpinning design

Mayes and De Freitas (2005) identify three broad perspectives in educational theory which make varying assumptions regarding understanding of the learning process. These perspectives are the associationist/empiricist perspective (learning as an activity, with behavioural objectives), the cognitive perspective (learning as achieving understanding), and the situative perspective (learning as social practice). The three views are now presented.

3.5.1.1. The associationist/empiricist perspective

In this approach, subject matter is analysed as specific associations and expressed as behavioural objectives. Based on a task analysis, learning tasks are arranged in sequences based on their relative complexity. Simpler activities are prerequisites for more complex tasks. Learning is described as the formation, strengthening and adjustment of associations, particularly through the reinforcement of particular connections through feedback. Implications for design are the individualisation of instruction, where each student responds actively to questions or problems, and the importance of providing immediate feedback. This is especially relevant to the development of programs for the teaching of routine skills.

3.5.1.2. The cognitive perspective

Knowledge acquisition and understanding are viewed as the outcomes of an interaction between new experiences and existing mental structures. This means that the learner's key cognitive challenge is the building of a structure for understanding the subject matter. As performance becomes more fluent, the component skills become automatised, meaning that conscious attention is no longer required for lower-level aspects of performance and this frees up cognitive resources for more complex levels of processing. This approach is in sharp contrast to the former view of learning as the strengthening of associations.

3.5.1.3. The situative perspective

This social perspective on learning views all learning as 'situated', since a learner will always be subjected to influences from the social and cultural setting, or context, in which the learning occurs. The learning outcomes will at least partly be defined by this situational setting and involve the abilities of learners to participate successfully in the practices of the communities in which the knowledge is situated. In this perspective the focus is not on analyses of subtasks, but on the patterns of successful practice.

Barab & Duffy (2000) defines an activity-based view of situated learning as *practice fields*, which represent constructivist tasks where the learning activity is represented as authentically as possible within the social context in which the skills or knowledge are normally embedded. The main design emphasis should be on the relationship between the nature of the learning task in the training environment, and its characteristics when situated in real use.

Regarding all three perspectives mentioned above, Mayes and De Freitas (2005:11) suggest that most implementations in e-learning will "include blended elements that emphasise all three levels: learning as behaviour, learning as the construction of knowledge and meaning, and learning as social practice".

3.5.2. Instructional design frameworks and multimedia learning

According to Chen (2010), traditional instructional design models offer no precise guidance for the process of designing instruction in virtual reality. Literature that describes how existing instructional design is being used to develop instruction in virtual reality environments remains scarce (Soto, 2013). Lau, Yen, Li and Wah (2014) indicate that virtual reality systems can be designed for learners to actively experience different situations and gain hands-on experience in problem solving, rather than simply discovering and perceiving information.

Since multimedia technologies broadly refer to the development and use of various types of media to enhance content visualisation and user interaction, instructional designers in virtual reality learning systems should apply multimedia learning principles (Lau *et al*., 2014). It is important to note, however, that even though the use of multimedia, such as virtual environments, can provide a richer learning experience, it does not guarantee effective learning. Clark and Taylor (1994) claim that learning that occurs due to exposure to media is caused by the instructional method embedded in the presentation, and not by the media. Clark and Mayer (2003:2) state unambiguously that "learning results from designing lesson materials with the right instructional methods, regardless of how the lesson will be delivered". Zhang, Wang, Zhao, Li and Lou (2008:155) concur when they state that "multimedia instruction messages that are designed in light of how the human mind works are more likely to lead to meaningful learning than those that are not".

A study by Liao (1999), in which 35 studies were analysed, concluded that multimedia-based instruction is superior to traditional instruction, but Zhang *et al.* (2008) point out that of the 35 studies investigated by Liao, ten of them actually indicated that traditional instruction was superior to multimedia instruction. The reason given for this inconsistency is poor instructional design of the multimedia instruction for those ten cases, where learner-centred principles were not followed and the designs did not correspond to learners' cognitive modes.

Brunken, Plass and Leutner (2004) state that many instructional design experiments have been conducted by educational technology researchers to determine how learners can benefit most from multimedia learning environments, but that the research focus has shifted to integrated models of cognitive processing. Examples of theoretical frameworks used in research on instructional design effects and individual differences in information processing, are cognitive load theory (Sweller, 1999) and the cognitive theory of multimedia learning (Mayer, 2002), both of which are considered in the next subsections.

3.5.2.1. Cognitive load theory

Cognitive load theory (CLT) states that there is a limit to the amount of cognitive capacity a learner can devote to a specific learning activity. This capacity is distributed over several cognitive processes required for learning and is described in three forms of cognitive load (Paas, Renkl & Sweller, 2003; Sweller, Ayres & Kalyuga, 2011).

- Intrinsic load: This depends on the material to be learned. The demand on cognitive capacity depends on the interrelationships between elements to be learned and their complexity.
- Extraneous load: The learning material can be presented in various ways using different instructional designs that could require varying amounts of cognitive capacity, independent of the learning content. The term *extraneous load* refers to cognitive capacity that may be required to compensate for a poor instructional design. Extraneous load does not contribute to the learning process.
- Germane load: This is the capacity required for the actual learning to take place, including the understanding of the new learning material, schema construction and integration, and storing of material in memory.

Since the total available cognitive capacity is limited, and the intrinsic load is assumed constant, it is argued that cognitive resources for germane activities can be freed up by minimising the extraneous load. Therefore, the instructional design of learning materials plays a major role in effective knowledge acquisition.

Extraneous load can be minimised by eliminating redundant and irrelevant elements, but Brunken *et al.* (2004) warn that this may also lead to design with a low level of interest. They suggest that reducing the extraneous load can be better achieved by taking into account the complex interaction between the presentation mode, the learning process and the learning material. More research is required in order for instructional designers to understand this complex interaction and be able to design interesting learning material without imposing too much extraneous load on the learner.

Chalmers (2003) named CLT principles for decreasing extraneous cognitive load as a means of increasing the usability of educational computer systems. Sawicka (2008) pointed out that designing usable learning environments reduces extraneous cognitive load and may contribute to improved learning. Similarly, Morrison, Dorn and Guzdial (2014) and Mason, Cooper and Wilks (2015) reached the same conclusion when assessing the cognitive load of lectures in introductory programming courses.

3.5.2.2. Cognitive theory of multimedia learning

Mayer and Moreno (2003) presented the cognitive theory of multimedia learning (CTML), which is based on three principles of learning from cognitive psychology.

- Dual channels: Humans have separate channels for processing information of a visual or pictorial nature and information that is auditory or verbal. For example, video images are processed in the visual channel and narrations are processed in the auditory channel.

- Limited capacity: Humans can only process a limited amount of information in each channel at any one time.
- Knowledge construction: Humans learn by mentally organising information in coherent structures and integrating it with prior knowledge.

As indicated in Figure 3.3, the CTML specifies five cognitive processes involved in multimedia learning as:

- selecting relevant words from a presentation,
- selecting relevant images from a presentation,
- organising the selected words into a coherent verbal representation,
- organising the selected images into a coherent pictorial representation, and
- integrating the verbal and pictorial representations with prior knowledge.

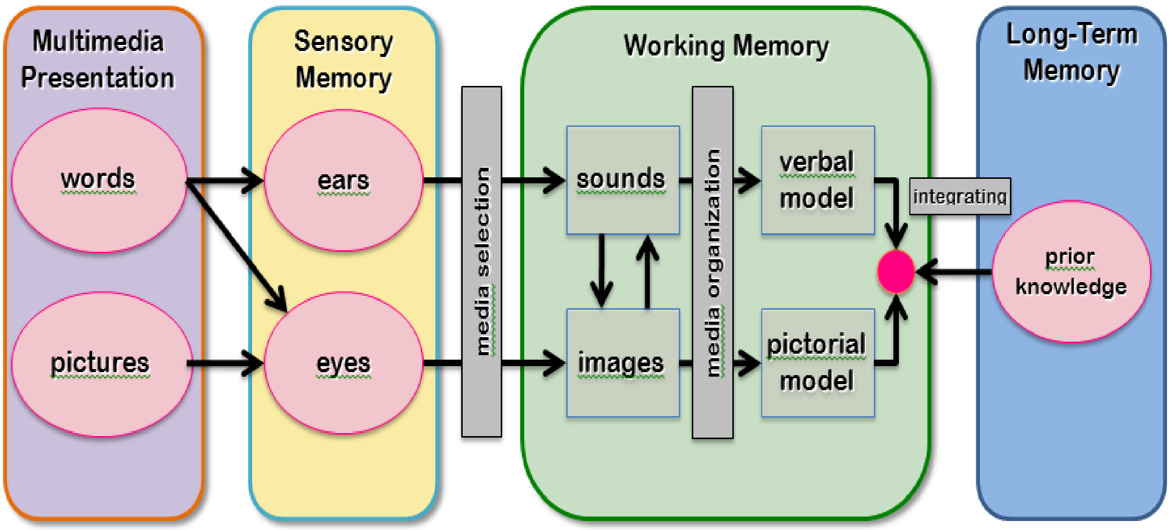


Figure 3.3: Schematic presentation of the cognitive theory of multimedia learning.

Similar to the comments made by Brunken *et al.* (2004) regarding CLT, Mayer, Fennell, Farmer and Campbell (2004) present two important design principles to improve multimedia learning: reduce the cognitive load to free up working memory; and increase learner interest, which will encourage them to use the freed memory capacity for processing subject matter during learning. One way of stimulating interest is to present the material in a visually appealing way.

Hede and Hede (2002) identified twelve conceptual elements that are inter-related and may affect the potential for learning with multimedia: visual input, auditory input, learning style, learner control, attention, working memory, cognitive engagement, learning, motivation, intelligence, reflection and long-term storage of information. In order to design effective learning tools, designers of multimedia learning products should be aware of the relationships among these elements. Mayer (2014) encourages instructional designers applying CTML to explore techniques for increasing learner motivation in the context of learning via multimedia.

3.5.2.3. Instructional design principles for multimedia

The use of multimedia offers designers opportunities to design meaningful and effective learning environments. This is not achieved simply by combining different media, but multimedia should be used mindfully in ways that augment the learning experience.

Based on literature studies and years of experience, Zhang *et al.* (2008) propose seven multimedia design principles.

- Usability Principle: Interfaces should be designed in ways that simplify the process of learning to use the software and that allow learners to focus on the educational material being presented. Usability patterns can provide important information to inform interaction design decisions.
- Multimodality Principle: Using auditory and visual working memory in combination can increase working memory capacity and understanding, if the information directed at each channel is integrated with information in the other channel. When words and pictures are both presented, learners have the chance to construct verbal and visual cognitive representations and integrate them (Chareen, 2007; Moreno & Mayer, 1999).
- Contiguity Principle: Learning is increased when corresponding narration and animation are presented simultaneously. Mayer and Moreno (2003) distinguish between temporal and spatial contiguity. Temporal contiguity means that corresponding words and pictures should be presented at the same time, while spatial contiguity means that corresponding words and pictures should be presented near, rather than far from each other on a page or screen.
- Coherence Principle: Better learning occurs when extraneous material such as irrelevant text, video or animation and superfluous graphics or sound are excluded (Moreno & Mayer, 2000).
- Redundancy Principle: The redundancy effect occurs when information that can be fully understood in isolation, as either visual or auditory information, is

presented via both channels and yet is essentially the identical information. This will increase working memory load, which interferes with the transfer of information to long-term memory.

- Pre-training Principle: Learning from a multimedia presentation will be more effective if learners are already familiar with the fundamental components of the presentation. Attempting to integrate components into a consolidated whole can rapidly overload working memory. Exposing learners to pre-training ensures that they already possess basic schemas of learning components, before they are exposed to material that requires them to integrate each component into larger schemas.
- Learner Control Principle: Learning is increased when the pace of presentation is controlled by the learner, rather than by the program. It is important to provide control mechanisms that can be customised to the abilities and styles of different learners (Hesham, 2004).

Canqun and Zhonghua (2010) present design principles that should be applied when designing multimedia courseware.

- Interactivity: Applying interactivity can make the process of accessing and using information more active and positive for the user, rather than a passive experience.
- Integrity: The content integrity of the knowledge components in non-linear courseware should be carefully considered.
- Consistency: This refers to not only consistency of the input and output, but also of the interface design of similar elements for analogous interfaces.
- Compatibility: There should be high compatibility between the interface designers and the expectations of the learners.
- Simplicity: Screen layouts and prompt information should be concise and clearly expressed.
- Sanity: High fault tolerance with simple and fast error correction should be implemented.

Chen, Toh and Fauzy (2004) proposed a theoretical framework for the instructional design of desktop VR learning applications. This framework was implemented in the development of an application to assist novice car drivers in comprehending traffic rules. It comprises a macro strategy and a micro strategy.

The macro strategy is a combination of the integrative goal concept of Gagne and Merrill (1990) and the learning model for constructivist learning proposed by Jonassen (1999).

The integrative goal approach includes individual objectives relating to verbal information, intellectual skills and cognitive strategies. Winn (1993), Jonassen (1999), and Kim, Park, Lee, Yuk and Lee (2001) all refer to VR's capabilities of affording constructivist learning. The macro strategy of Chen *et al.*'s (2006) framework includes components for problem context, problem presentation and problem manipulation, which encourage the design of constructivist learning environments where learners learn through their efforts to solve problems.

The micro strategy is based on the cognitive theory of multimedia learning (Mayer, 2002). This theory assumes that the human information processing system includes two channels: a visual or pictorial channel and a channel for auditory or verbal processing. Each channel has limited processing capacity, and active learning entails carrying out a coordinated set of cognitive processes during learning. Through a series of experiments Mayer (2002) came to the following conclusions:

- When words and pictures are both presented, learners have an opportunity to construct verbal and visual mental models and build connections between them.
- When corresponding words and pictures are near to each other on the screen, learners do not have to use cognitive resources to visually search the page or screen and learners are more likely to be able to hold them both in working memory at the same time.
- Extraneous material competes for cognitive resources in working memory and can divert attention from the important material, can disrupt the process of organising the material, and can prime the learner to organise the material around an inappropriate theme.
- When pictures and words are both presented visually, the visual channel can be overloaded while the verbal channel is unused. When words are presented auditorily, they can be processed in the verbal channel, thereby leaving the visual channel to process only the pictures.

Mayer (2008) suggests principles that summarise and consolidate the essence of good multimedia design. These evidence-based and theoretically grounded principles are further explained by Clark and Mayer (2011). They are categorised into principles for reducing extraneous processing; principles for managing essential processing; and principles for fostering generative processing, and are outlined in Table 3.2.

Table 3.2: Principles for multimedia learning (Mayer, 2008; Clark & Mayer, 2011).

Principles for reducing extraneous processing	
Principle	*Definition*
Coherence	Reduce extraneous material
Signalling	Highlight essential material
Redundancy	Do not add on-screen text to narrated animation
Spatial contiguity	Place printed words next to corresponding graphics
Temporal contiguity	Present associated narration and animation simultaneously
Principles for managing essential processing	
Principle	*Definition*
Segmenting	Present animation in learner-paced segments
Pre-training	Provide pre-training in the characteristics of key components
Modality	Present words as spoken text rather than as printed text
Principles for fostering generative processing	
Principle	*Definition*
Multimedia	Present words and pictures rather than words alone
Personalisation	Present words in conversational style rather than formal style

Alessi and Trollip (2001) indicate four phases that constitute a generic instructional model for effective and efficient learning.

- Presentation of information: To teach new learning content, information must first be presented, for example, by way of rules or examples. Using examples is a good way of initially presenting information, though more than one example may be necessary before learners gain the required skill or apply the rules. Even though the presentation of information is a basic behaviourist principle, it is also supported by constructivists. Duffy and Jonassen (1991) advise that learners should be provided with some explicitly expressed knowledge as a starting point to the learning process. This can be done as an instructor-centred or media-centred activity.
- Guiding the learner: This stage is more interactive and involves both the learner and the medium. The role of the educator is to observe the learner, correct errors

and give suggestions or hints. The most common method of guiding learners is asking them questions and providing feedback to their responses.

- Practice: Repeated practice is required to ensure learning, and should result in speed, fluency and retention. Practice is a learner-centred activity that emphasises the learner practising, while the instructor makes brief supportive or corrective statements.
- Assessing learning: It is important to evaluate to what extent learning has occurred. According to Alessi and Trollip (2001), evaluation, termed *assessment* in South Africa, should be done not just to grade learners, but also to guide instructional decisions, such as determining the varying instructional needs of different learners.

3.5.3. Methodologies for interactive multimedia that facilitate learning

It is important to note that in the instructional design of e-learning and e-training, a combination of forms or methodologies is likely to be used (Alessi & Trollip, 2001; De Villiers, 2005a). Each of the four phases above can be facilitated by one or more of the various forms or methodologies that use interactive multimedia.

- Tutorials: These are normally used for the first two phases of instruction. Information is presented and learners are guided in their first encounters with the information. Interactive dialogue is typically used when coaching learners, alternating presentation with questioning, and providing feedback on their responses.
- Hypermedia: These programmes are less structured than tutorials and are often used in constructivist or open-ended learning experiences. They do not have to follow a linear style but can be sequenced flexibly using hyper-links, branches and networks. This gives learners some individual control as they choose their own paths.
- Web-based learning environments: This medium can be combined with any other methodologies for any of the four phases of instruction. It is a delivery medium currently generally used together with hypermedia.
- Drills: These are designed to encourage practice to enable learners to become fluent and to retain information. Drills are often used together with educational games for motivation.
- Simulations: In any of the four phases of instruction, simulations can be used to simulate real-world situations. When combined with games, discovery learning is fostered.
- Games: Games are frequently used for practice, and can be combined with other

methodologies as mentioned above, with drills to make them less tedious, with simulations to create discovery environments. Performance can be extrinsically motivated through the use of games.

- Tools and open-ended learning environments: Computer applications such as databases and spreadsheets are tools for cognitive computing activities that can be controlled by the learners themselves and used together with other media. Being more flexible and open-ended, they support constructivist learning and can be used during any of the phases of instruction.
- Tests: Normally in the final phase of instruction, tests are used to summatively assess learners. However, as part of formative assessment, practice tests and quizzes are useful for interactive practice.

3.6. Usability evaluation

Usability evaluation is concerned with gathering information about the usability or potential usability of a system in order either to improve its interface or to assess it. Evaluation is needed to check that users can use the product and that they like it, particularly if the design concept is new (Rogers *et al.*, 2011). Evaluation is concerned with gathering data about the usability of a design or product by a specified group of users for a particular activity within a specified environment or work context (Preece, Rogers, Sharp, Benyon, Holland & Carey, 1994). According to Rogers *et al.* (2011), the three main evaluation approaches for interactive systems are usability testing, field studies and analytical evaluation.

Various usability evaluation methods (UEMs) exist and several authors offer classifications of the techniques used to measure usability-related factors (Avouris, 2001; Dix, Finlay, Abowd & Beale, 2004; Preece, Rogers & Sharp, 2002; Shneiderman & Plaisant, 2005). This section will discuss UEMs by categorising various methods according to the two classes defined by Foltz, Schneider, Kausch, Wolf, Schlick and Luczak (2008), who separated evaluation methods into empirical evaluation methods and analytic evaluation methods. Empirical evaluation methods involve actual or designated users. The methods can be relatively informal, such as observing people while they explore a prototype, or they can be quite formal and systematic, such as a tightly controlled laboratory study or a comprehensive survey of many users. Analytic evaluation methods involve expert analysis and can be used early in the system development process, before there are users or prototypes available for empirical tests.

3.6.1. Empirical evaluation methods

The empirical, or experimental, evaluation approach to evaluation originates from the scientific and engineering fields where experiments have reliably been used for precise measurement of issues. Empirical evaluation is based on the use of scientific experimental methods to test hypotheses about the usability of a system (Preece, 1993).

According to Shneiderman (1998), traditional experimental methods have been found to be used by researchers in the HCI discipline in the study of computer system interfaces. When an experiment is carried out, the purpose is to answer a question or test a hypothesis. In this way new knowledge is discovered. The method usually determines the relationship between variables by manipulating one of them and observing what effect it has on the others (Preece *et al.*, 2002). The researcher, within the context of HCI, manipulates factors associated with the interface of the computer system. The effect of this on aspects of user performance is then studied (Preece, 1993:117).

Despite its reliability, empirical evaluation as an evaluation strategy in HCI is not frequently used as it can be very expensive and may require sophisticated equipment (Ardito, Costabile, De Marsico, Lanzilotti, Levialdi, Roselli & Rossano, 2006).

The next subsections discuss three empirical evaluation methods: observational methods, usability testing, and query techniques which include interviews and questionnaires.

3.6.1.1. Observational Methods

Observational methods of usability evaluations are performed by observing the actual users interacting with the system. The users can be observed in their natural setting, or when performing a set of predetermined tasks in laboratory-like conditions (Dix *et al.*, 2004; Preece *et al.*, 2002). Rogers *et al.* (2011) define observation as a useful data-gathering technique at any stage during product development. The advantage of observational methods is that the usability problems of real users can be directly identified. According to Dix *et al.* (2004), the two main observational techniques used in usability evaluation of computer systems are think-aloud and protocol analysis.

During a think-aloud evaluation, users are not only observed but also asked to elaborate verbally on their actions by describing what they are trying to do, their perceptions of what is happening, and why they chose a particular action (Dix *et al.*, 2004). During this process, the evaluator's role is to be supportive, providing prompts and listening for clues about the usability of the system, being careful not to give instructions or cause any distractions. After completion of the evaluation session, the evaluator would invite the participant to make any comments or suggestions, or answer any questions.

Apart from being simple to use since little expertise is required, think-aloud also provides a good understanding of the user’s mental model and interaction with the system. Possible concerns regarding this technique are that users may change their behaviour when they are aware of being observed or that the process of verbalising may distract users from performing in the way they normally do (Preece, 1993; Rogers *et al.*, 2011).

Protocol analysis refers to the analysis that takes place after the observation and is based on the evaluator’s record of occurrences during the evaluation session. This record is called a protocol. Dix *et al.* (2004) describe different methods that can be used for recording user actions. These include paper and pencil, audio recording, video recording, and computer logging.

3.6.1.2. Usability testing

Usability testing is a formal, controlled observational technique that involves measuring the performance of users as they undertake tasks in an interactive system. It has been shown to be an effective method to improve usability and to rapidly identify potential problems (Dumas, 2003; Dumas & Reddish, 1999). Usability testing is conducted by usability specialists in laboratories that are usually equipped with sophisticated observation equipment such as audio-visual recording facilities and one-way glass.

Specialised equipment can be used to conduct eye tracking and to gather physiological measurements. The equipment is expensive, specialist skills are required, and the process is time-consuming (Moczarny, De Villiers & Van Biljon, 2012). Monitoring of physiological responses such as is done in eye tracking, is associated with usability testing since it is usually conducted in an HCI laboratory. Pool and Ball (2006) defined eye tracking as a method to determine eye movement and eye-fixation patterns. Measuring not only where people look, but also their patterns of eye movement, may indicate to the tester which areas of a screen users find easy or difficult to understand (Dix *et al.* 2004).

3.6.1.3. Query techniques

Query techniques are relatively simple and inexpensive to administer and they support the philosophy that the best way to identify a system's usability problems is to ask the user directly (Ardito *et al.*, 2004a; Dix *et al.*, 2004). The two main query techniques, interviews and questionnaires, are well established in HCI research.

Interviews

Interviews are ways to gather information directly from individual users. The interviewer verbally asks about the usability of the system, and is able to focus on particular issues of concern. This can lead to helpful and constructive suggestions (Shneiderman & Plaisant, 2005). An advantage of interviews, as indicated by Dix *et al.* (2004), is the possibility of varying the level of questioning depending on the particular context. An interviewer can start with a general question about a task before progressing to specific questions, and probing more deeply, if necessary, as different issues arise.

According to Genise (2002), using interviews for evaluation has advantages in that they are useful for obtaining detailed information, only a few participants may be required, and interviews serve well when conducted after some other UEM so as to follow up on issues that emerged. Four types of interviews can be used depending on the evaluation goals (Preece *et al.*, 2002).

- Unstructured interviews: An unstructured, or open-ended interview uses questions that allow for the interviewee to freely express his/her own opinion. There is no predetermined direction but the interviewer must ensure the interview is within the scope of prescribed goals. A potential major benefit of the unstructured interview is that interviewees may mention facts not anticipated by the interviewer, which can be probed further. A disadvantage is that a great deal of unstructured data may be generated which can be very time-consuming and difficult to analyse.
- Structured interviews: The interviewee is asked a set of predetermined questions, according to a fixed protocol which does not vary from one interviewee to another, making a structured interview easier to conduct. It is also easier to analyse the data from such interviews. Structured interviews should be used when specific questions can be identified in line with a clear understanding of the goals of a study (Preece *et al.,* 2002).
- Semi-structured interviews: Both closed and open questions can be used. For consistency, the interviewer has a basic script to ensure that the same topics are covered with each interviewee. The semi-structured interview will usually

commence with a set of pre-planned key questions, and then an opportunity is given to the interviewee to elaborate or provide more relevant information. Furthermore, interesting and unanticipated areas can be probed further.

- Group interviews: Shneiderman & Plaisant (2005) emphasise the use of focus-group discussions or group interviews following a series of individual interviews. This is done to collaboratively explore the general nature of the comments from different individuals.

Questionnaires

Questionnaires are a thoroughly established technique for collecting demographic data and users' opinions (Preece *et al.*, 2002). Open questions would allow a respondent to freely express his/her own answer, whereas closed questions provide a choice of options from which to choose.

Thorough preparation is vital before a major survey is carried out, the questionnaire reviewed and then pilot-tested with a small group to avoid potential misunderstandings. Shneiderman & Plaisant (2005) advise that the design for the statistical analysis and presentation of data should also be planned beforehand. Preece *et al.* (2002) have set out guidelines to assist in the preparation of questionnaires:

- Questions should be clear and specific.
- Where possible, ask closed questions with a range of answers to choose from.
- Questions asking for an opinion should offer an option for a neutral opinion.
- Carefully consider the order of questions as sequence can influence responses.
- Avoid jargon and consider whether different questionnaires will be needed for different populations.
- Give clear instructions on how the questionnaire is to be completed.
- Long questionnaires may deter participation, so balance the use of white space with the need to keep the document compact.
- When scales are used, make sure the ranking is intuitive and consistent. For example, in a Likert scale of 1 to 5, 1 should indicate low agreement and 5 should indicate high agreement consistently throughout the questionnaire.

User satisfaction questionnaires are survey instruments which are administered to participants, with little input from the researcher. The questionnaire has a structured format so as to obtain uniform data from participants. A large number of responses may be gathered and analysed effectively and efficiently (Shneiderman and Plaisant, 2005).

Due to questions being predetermined and fixed for all users, and not customised to individuals, questionnaires may be less flexible in comparison to some other methods, such as interviews. The advantages of using questionnaires are that they can reach a wider subject group as compared to interviews, and are inexpensive and easy to use (Dix *et al.*, 2004).

The advantages of using interviews and questionnaires make query techniques attractive for evaluation of the VR training systems proposed in this study. These techniques are applied in this study as described in Chapter Five.

3.6.2. Analytic evaluation methods

In an analytic evaluation method, experts inspect the human computer interface so as to predict problems users would face when interacting with it. This method is an alternative to conducting evaluation with end-users, and was introduced after recognising that users are not always easily accessible or that involving them would make the evaluation process too expensive or time consuming. The two main expert evaluation techniques are heuristic evaluation and walkthroughs. Rogers *et al.* (2011) point out that these techniques are generally easy to learn, inexpensive and effective in identifying usability problems. The next subsections discuss three types of analytic evaluation methods: model-based evaluation methods, walkthroughs and heuristic evaluation.

3.6.2.1. Model-Based Evaluation Methods

In model-based evaluation, a model of how users would use a proposed system is used, and predicted usability measures are obtained by calculation or simulation (Kieras, 2003). In terms of the physical and cognitive operations that must be performed by the system, system designers are able to make use of this approach to analyse and predict expert performance of error-free tasks. According to Rogers *et al*. (2011), these methods are suitable for usability evaluation in an early phase of system development.

The GOMS model (goals, operations, methods and selection) is particularly useful in helping make decisions on the effectiveness of new products. Because the GOMS model gives quantitative measures of user performance, it allows for comparative analysis to be performed for different prototypes, interface or specifications relatively easily. The limitations of GOMS are that it is intended to be used only to predict expert performance and does not allow for errors to be modelled, and that it can only model small sets of routine computer-based tasks (Hochstein, 2002). These limitations make model-based evaluation unsuitable for general use in usability evaluations of interactive e-learning applications such as VR training systems.

3.6.2.2. Walkthroughs

Walkthroughs, like heuristic evaluation, are methods of predicting users' problems without doing user testing. They are carried out by experts 'walking through' the tasks and recording the problematic usability features. Cognitive walkthroughs do not involve users, but pluralist walkthroughs involve a team made up of users, developers and usability experts (Rogers *et al.*, 2011; Madan & Dubey, 2012).

Cognitive walkthrough

Cognitive walkthrough involves evaluators working through a sequence of steps likely to be followed by users when they interact with the system. Expert evaluators do a mental stepthrough of expected actions of users, providing insight into the ease with which the system is learned and used (Dix et al., 2004; Shneiderman and Plaisant, 2005). The main function of a cognitive walkthrough is to determine how easy it is to learn to use the system.

After completion of the walkthrough, a usability problem report is written about what works well and what needs to be improved. The main advantage of cognitive evaluation is that it focuses on users' problems in detail without involving users themselves, but it is very time-consuming and not easy to perform (Rogers *et al.*, 2011).

Pluralistic walkthrough

In this evaluation method, a group of users, developers and usability experts collaboratively step through a set of tasks to discuss and evaluate the usability issues associated with the system, with all participants assuming the role of the user during the walkthrough. Benefits of pluralistic walkthroughs are that they can be used in the early stages of system development and feedback can be given to the developers about the design of the system, and there is increased buy-in by the users, since they participate in the development of the system (Bias, 1994; Hollingsed & Novick, 2007). A major disadvantage is that the moderator determines which path in the system to follow, with the result that not all the possible paths are explored (Gulati & Dubey, 2012).

3.6.2.3. Heuristic evaluation

Heuristic evaluation is an inspection technique whereby experts apply a set of usability principles called *heuristics*, to evaluate whether a user interface conforms to these principles (Hix & Gabbard, 2002; Madan & Dubey, 2012; Zaibon & Shiratuddin, 2010). A heuristic is defined by Dix *et al.* (2004) as a guideline or general principle used to guide a design decision or to critique a decision that has already been made. It can also be referred to as a criterion. Paddison and Englefield (2003) describe heuristic evaluation as one of the most established and cost-effective techniques for usability evaluation of systems.

Heuristic evaluation involves having a small set of evaluators who examine the interface of the system and judge its compliance with the heuristics. Nielsen (1994) recommended that three to five evaluators be used since not much additional information is gained using a larger number. If the evaluators are experts in both HCI and also in the domain area of the application, Karoulis and Pombortsis (2003) indicate that usually two to three of them will point out the same percentage as three to five HCI experts. Results are even higher if the expert evaluators have dual domain backgrounds, such as expertise in both usability and the topic/domain at hand (Georgson, Weir & Staggers, 2014). Heuristic evaluation is appropriate to various development phases, from initial prototype to early design evaluations (Nielsen, 2005; Shneiderman & Plaisant, 2005; Ssemugabi & De Villiers, 2010).

Some of the advantages of using heuristic evaluation include the following:

- It is an informal evaluation method that is relatively effective, inexpensive, and easy to perform (Karoulis & Pombortsis, 2003; Ardito *et al.*, 2006; De Kock, Van Biljon & Pretorius, 2009).
- It can result in major improvements to a particular user interface (Belkhiter, Boulet, Baffoun & Dupuis, 2003; Karoulis & Pombortsis, 2003).
- During a short session, a small number of experts can identify a range of usability problems.
- Experienced evaluators can suggest solutions to usability problems that individual users may not pick up.

Two decades ago, Preece (1993) identified certain disadvantages of heuristic evaluation that are equally valid today:

- Experts may be biased due to their strong subjective views and preferences, and this may lead to biased reports.

- It may be difficult to find evaluators who are experienced in both the specific domain of the system and in HCI research.
- Evaluators may require a great deal of information about the knowledge level of the users, their typical tasks and their responses to problems.
- Expert evaluation may not capture the variety of real users' behaviours. For example, novice users may perform unexpected actions that an evaluator might not think of.

In this study, heuristics were developed to evaluate the training prototypes developed for the mining industry, and six dual domain experts were used for the evaluation.

3.6.3. A synthesised framework for heuristic evaluation of virtual reality training applications

Various factors impact on determining suitable theoretical foundations for e-learning applications. No single paradigm is appropriate for all situations, since domain, context and content all have to be considered (De Villiers, 2005a). Technological issues and underlying educational theories should be considered, as well as usability, in order to find an appropriate solution. Ardito *et al.* (2006) point out that there should be a synergy between the learning process and the interaction with an e-learning application. Costabile *et al.* (2005) advise that evaluations of educational software should investigate both pedagogical effectiveness and usability aspects. In the integration of usability and learning, usability features that are important for the achievement of educational goals should be addressed (Squires & Preece, 1996).

Ardito *et al.* (2006), advocate that specific custom-designed guidelines should be provided for the evaluation of e-learning, rather than using a small set of general criteria. In line with this call, this study synthesises a new set of guidelines specifically customised for evaluation of VR training applications within the specific context of mining safety. The synthesised evaluation framework is presented in Section 5.8 in the chapter on research design.

3.7. Conclusion

This chapter overviewed user-centred systems design, learner-centred design, interaction design and usability engineering as methodologies of user-centred design. An important emphasis was placed on usability, specifically usability of virtual environments and e-learning. Because e-training applications, as proposed by this study, should be based on sound learning theories, three main theories of learning were discussed: behaviourism, cognitivism and constructivism. Since learning is facilitated by reinforcing correct performances, the behaviourist principles of reinforcement, retention and transfer of learning are important design considerations. The instructional designer needs to provide reinforcing activities to promote retention of the learning material. Practice provides increased opportunities for reward, reinforcement and the creation of cognitive structures, which lead to more efficient use of long-term memory.

Instructional design aspects of multimedia learning were also addressed, including cognitive load theory, the cognitive theory of multimedia learning, instructional design principles for multimedia, and methodologies for interactive multimedia that facilitate learning. Cognitive load theory states that working memory can process only a few elements at any one time. Four of the principles of multimedia learning (presented in Table 3.2) – contiguity, coherence, modality and redundancy – also reflect the theme that learning is improved when working memory is not overloaded. Furthermore, constructivist learning is more likely to occur when learners have corresponding visual and verbal representations in working memory at the same time.

The last part of the chapter explained various usability evaluation methods, categorised into empirical evaluation methods and analytic evaluation methods, with a focus on heuristic evaluation. Their application in the present research was mentioned.

In Chapter Five, a synthesised evaluation framework is proposed for the heuristic evaluation of VR training applications. The framework comprises four categories of criteria: Instructional Design, General Usability, Virtual Reality Systems Design and Context-specific heuristics. The present chapter and others contribute to the theoretical foundations of this evaluation framework. Instructional design and usability aspects relevant to e-training were discussed in Sections 3.5 and 3.3 respectively. Aspects related to VR systems design are detailed in Chapter Seven (Section 7.3), while the context of safety training in mining is discussed in Chapters Four (Sections 4.2 and 4.3) and Six (Section 6.5).

The topics covered in this chapter assisted the researcher in synthesising the evaluation framework, which is the main deliverable of this study. The evaluation framework will be applied to evaluate two prototype systems for mine safety training. To improve understanding of the context in which safety training occurs, the next chapter, Chapter Four, discusses mine safety practice.

Chapter Four
MINE SAFETY PRACTICE

4.1. Introduction

This literature study focuses on current safety practices in the South African mining industry. Despite stringent safety legislation, accidents in the mining industry are still causing high numbers of fatalities and injuries.

This chapter moves beyond the literature and lays the foundations for the empirical work of this research, by relating the theory and legislation to application. Section 4.1 briefly introduces the South African mining environment. This is followed by an overview in Section 4.2 of the safety legislation applicable to the mining industry. A discussion of the major stakeholders involved in the industry is the topic of Section 4.3, while mine safety statistics are presented in Section 4.4. Industry competency certificates are discussed in Section 4.5.

In 2007, the then State President, Thabo Mbeki, requested a mine health and safety audit to determine whether the mines meet health and safety standards as prescribed by the Mine Health and Safety Act. Section 4.6 deals with the findings of the presidential audit relevant to this study. Each major section in the chapter ends with a subsection called *Application to training*. These subsections link the topics discussed to the main focus of the study, namely the improvement of mine safety training using virtual reality. Section 4.7 concludes the chapter.

Figure 4.1 shows the layout of Chapter Four.

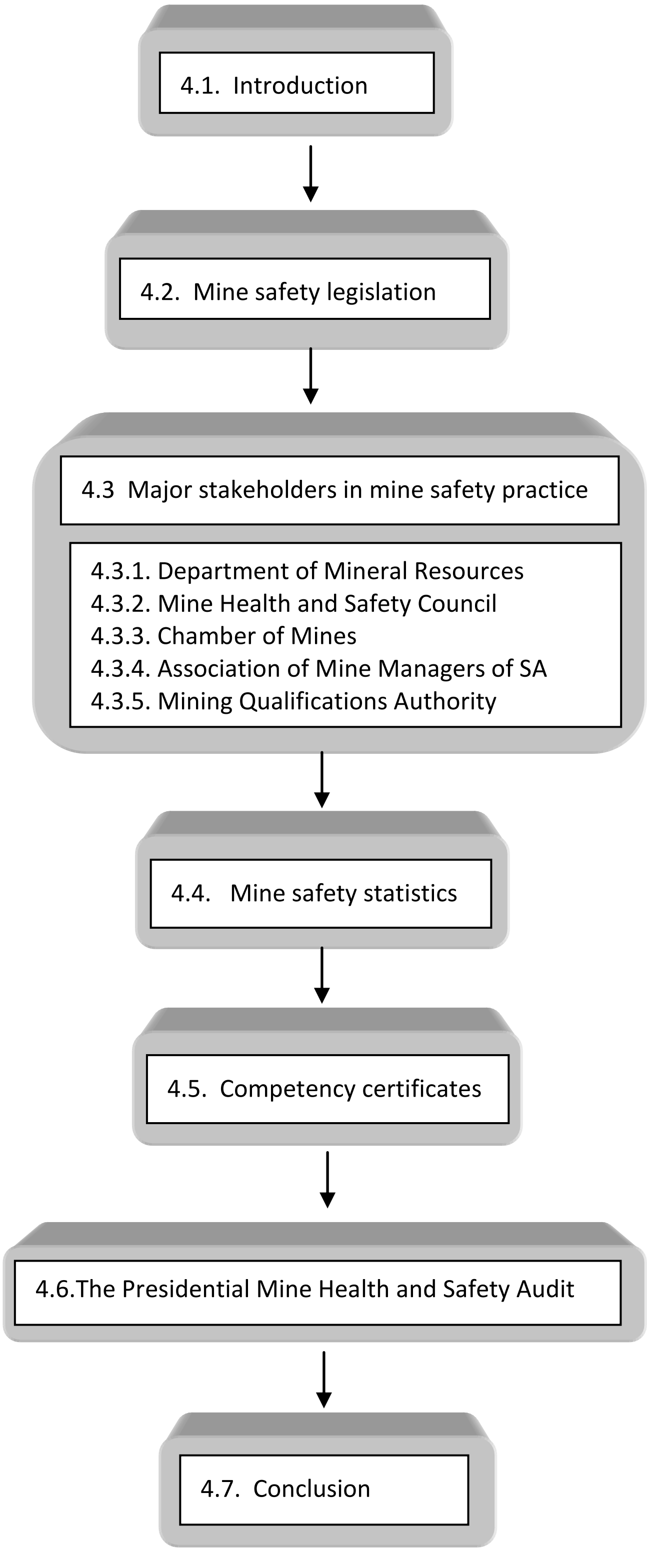


Figure 4.1: Layout of Chapter Four.

The South African mineral sales totalled R363.8 billion in 2012, and the mining sector accounted for 8.3% of the gross domestic product (GDP) directly, while the indirect contribution to GDP was estimated at close to 17% (CoM, 2013). South Africa holds the world's largest reported reserves of gold, platinum group metals, chrome ore and manganese ore, and is ranked as the world's richest country in terms of its mineral reserves, worth an estimated $2.5 trillion (NPC, 2013). As indicated in Table 4.1, South Africa is ranked amongst the world's top five producers of a number of commodities (DMR, 2013). South Africa is also home to the world's three biggest platinum producers – Impala Platinum Holdings, Anglo Platinum and Lonmin – as well as three of the world's top six gold producers: AngloGold, Harmony and Gold Fields. Most of the coal used in European power plants originates from South Africa (Sguazzin & Lourens, 2007).

Table 4.1: South Africa's world ranking for minerals production (DMR, 2013).

MINERAL	WORLD RANKING
Gold	2
Platinum	1
Coal	5
Chrome ore	1
Diamonds	5
Vermiculite	1
Vanadium	1
Manganese ore	1
Fluorspar	4
Titanium	2

Gold and platinum are mined at great depths, using labour-intensive underground 'conventional' drill and blast techniques. The drill and blast method in mining predominantly entails the use of hand-held pneumatic drilling machines to drill holes into the rock in what is called the mine face area. The rock is charged with explosives and then blasted. The blasted rock is then removed from the working places by means of scraper winches and transported by conveyor belts or rail. Locomotives pull hoppers to the shaft stations, from where the rock is hoisted to the surface. The exposed hanging wall is then supported using various forms of temporary or permanent support, for example, stick support, solid timber packs, grouted packs, roof bolts or hydraulic props.

Due to South African mining activities being more labour intensive than in countries such as Canada, Australia, Germany and the USA, a greater number of South African workers are exposed to health and safety risks than their counterparts in mining establishments in those countries. Table 4.2 shows the labour distribution between the different major commodities.

Close to 500 000 people work in the mining industry in South Africa, with the platinum and gold sectors employing the most workers. The empirical research of this study was done in the platinum sector.

Table 4.2: Labour numbers at major mining sectors, 2012 (DMR, 2013).

Mine Type	Persons	% Contribution
Gold	134 256	26.9%
Platinum	182 026	36.4%
Coal	75 989	15.2%
Diamonds	12 120	2.4%
Copper	3 250	0.7%
Chrome	16 202	3.2%
Iron ore	23 200	4.6%
Manganese	7 166	1.4%
Other	45 574	9.1%
All mines	499 783	100%

4.2. Mine safety legislation

The minerals industry faces challenges on a daily basis, such as hazardous geological conditions, as well as mining at depth and seismicity. Several legislative stipulations have been adopted in order to deal with the occupational health and safety challenges facing the industry. According to Schreiber and Kielblock (2004), many aspects of the South African health and safety regulatory system for mining could well be emulated by counterparts elsewhere in the world. This system includes the Mine Health and Safety Act of 1996, the Mine Health and Safety Amendment Act of 1997, the Mineral and Petroleum Resources Development Act of 2002, and the Mine Health and Safety Amendment Act of 2008.

The act that currently covers the management and regulation of mines is the Mine Health and Safety Act, 1996 (Act No. 29 of 1996), referred to as the MHSA (MHSA, 1996). The MHSA replaced the Minerals Act and Regulations of 1991, incorporating all the regulations of the 1991 Act (Adams, du Plessis, Gumbie & Willis, 2007). The regulations in the MHSA not only protect the interests of mine employees, but also contractors and visitors to mine sites.

Similar to the health and safety regulations of the International Labour Organisation, the MHSA also incorporates the duties and rights of the state, employers and employees. Fundamental principles emanate from this conception.

- The state: The state has the duty and responsibility to effectively regulate health and safety conditions at workplaces.
- The employer: Employers have a duty and responsibility to protect the health and safety of their employees who may be exposed to occupational health and safety hazards and risks. As far as is 'reasonably practicable', employers are required to provide and maintain a safe working environment. 'Reasonably practicable' is defined as relating to
  - the severity and scope of the hazard or risk concerned,
  - the state of knowledge reasonably available concerning the hazard or risk,
  - the availability and suitability of means to remove or mitigate the hazard or risk, and
  - the costs and benefits of removing or mitigating the hazard or risk.
- The workforce: The employee has the right to refuse to work in an environment that may endanger his/her health or safety.

Other stated purposes of the MHSA are as follows:

- Promote a culture of health and safety.
- Promote training and human resources development.
- Provide for the enforcement of health and safety measures.
- Provide for effective monitoring systems, inspections, investigations and inquiries to improve health and safety.
- Provide for appropriate systems of employee, employer and state participation in health and safety matters by establishing representative tripartite institutions to review legislation, promote health, and enhance properly targeted research (DMR, 2010a).

The MHSA of 1996 was amended through the Mine Health and Safety Amendment Act of 1997 and the Mine Health and Safety Amendment Act of 2008. The 1997 amendment introduced the concepts of risk assessment and occupational health and safety management to the mining industry and also reduced inconsistencies in the interpretation of the mining inspectorate's enforcement responsibilities. This process assisted in achieving a clear and consistent approach to the enforcement of the Mine Health and Safety Act. The 2008 amendment was aimed at addressing gaps that had been observed over the previous 12 years and amended the MHSA so as to review the enforcement provision, simplify the enforcement system, tighten offences, and strengthen penalties. Ambiguities in certain definitions and expressions were also removed, and certain amendments were necessary to ensure consistency with other laws, particularly the Mineral and Petroleum Resources Development Act of 2002 (McKay, 2009; Pressly, 2008).

The MHSA was further amended in 2010. The financial penalty which can now be imposed for a breach of safety procedures is an administrative fine of R1 million per incident. Moreover, appeals cannot be lodged against these new fines, and a review application challenging a fine would not suspend its implementation. In addition, criminal prosecution of persons who contributed to the accident may follow (Badenhorst, 2011).

Application to training

Since the advent of democracy, mining legislation in South Africa has undergone far-reaching changes in response to political developments in the country and the persistently high incidence of mining accidents. An aspect of such change is a call for the provision of improved training directly relevant to the South African mining environment, which has subsequently been backed by legislative changes.

The process of overhauling health and safety legislation began in 1995 with the Leon Commission of Enquiry into health and safety in the mining industry (Barry, 1995). This was initiated largely at the insistence of the National Union of Mineworkers (NUM). The Leon Commission contended that mine employers had failed to provide proper and effective training to employees, thereby failing to combat the alarming rate of deaths and reportable injuries, particularly from rockbursts and rockfalls. The Commission recommended that existing health and safety training at all levels in the industry be improved. Many of the recommendations made by this commission were subsequently incorporated into the MHSA of 1996.

An extract follows from Section 10 of the MHSA (DMR, 2010a). Specific aspects pertinent to this research are highlighted in bold font.

"Manager to provide health and safety training

10.(1) As far as reasonably practicable, every manager must -

> (a) provide employees with any **information, instruction, training or supervision** that is necessary to enable them **to perform their work safely** and without risk to health; and

(b) ensure that every employee becomes **familiar with work-related hazards and risks** and the measures that must be taken to **eliminate, control and minimise** those hazards and risks.

(2) As far as reasonably practicable, every manager must ensure that every employee is **properly trained** -

(a) to deal with every risk to the employee's health or safety that -

(i) is associated with any work that the employee has to perform; and

(ii) has been recorded in terms of section 11;

(b) in the measures necessary to eliminate, control and minimise those risks to health and safety;

(c) in the procedures to be followed to perform that employee's work; and

(d) in relevant emergency procedures.

(3) In respect of every employee, the provisions of subsection (2) must be complied with

(a) **before the employee first starts work**;

(b) **at intervals determined by the manager** after consulting the health and safety committee;

(c) before significant changes are introduced to procedures, mining and ventilation layouts, mining methods, plant or equipment and material; and

(d) before significant changes are made to the nature of that employee's occupation or work".

These developments in legislation in 1996 had significant implications for employers who must provide effective safety training of workers as a basic requirement (Squelch, 1998). From the employers' perspective, Davids (1997) suggested that if miners were better educated and trained, it should contribute to reducing accidents on mines. A study published in the American Journal of Public Health investigated what the effect had been of mining safety training legislation in the USA. The study concluded that a federal policy that required miners to undergo safety training had reduced the incidence of permanently disabling injuries (Monforton & Windsor, 2010).

However, to date, conventional training has not been able to reduce the incidence of accidents in the local mining industry significantly, hence the increased interest in using VR training through which to achieve this goal.

4.3. Major stakeholders in mine safety practice

A broad variety of organisations and associations are involved in the South African mining industry. They range from government departments and statutory councils to specific-interest groups, such as the South African National Institute of Rock Engineering (SANIRE) and the Mine Ventilation Society of South Africa (MVSSA). This section discusses the roles of some of the major stakeholders involved in mine safety.

4.3.1. Department of Mineral Resources

The Department of Mineral Resources (DMR) was formed in 2010 when the Department of Minerals and Energy (DME) was subdivided into two separate government departments: the Department of Mineral Resources and the Department of Energy. The DMR, through the Mine Health and Safety Inspectorate, has established national policy and legislation, as well as systems to monitor, audit and inspect mines (DMR, 2010b). The DMR also supports training in the mining industry and contributes to the development of qualifications, skills programmes and learnerships, and provides technical advice to the mining sector (AMMSA, 2012).

At the DMR, the Mine Health and Safety Inspectorate (MHSI), which was established in terms of the Mine Health and Safety Act of 1996, is responsible for safeguarding the health and safety of people working at mines or affected by mining activities. The MHSI is headed by the Chief Inspector, who is also chairperson of the boards of the Mine Health and Safety Council (MHSC) and the Mining Qualifications Authority (MQA).

The DMR has offices in all nine provinces and mine inspectors operate from these offices in conducting regular inspections and audits at mine workplaces. Differing approaches are employed by inspectors to secure compliance with health and safety standards and thus to improve the work environment for mine workers. These approaches range from advice on mining operations, to improvement notices to non-compliant employers. In extreme cases that pose immediate danger to employees, operations at workplaces are stopped so that corrective actions can be taken before work is resumed. Section 54 of the MHSA sets out an inspector's power to deal with dangerous conditions:

"(1) If an inspector has reason to believe that any occurrence, practice or condition at a mine endangers or may endanger the health or safety of any person at the mine, the inspector may give any instruction necessary to protect the health or safety of persons at the mine, including but not limited to an instruction that –

(a) operations at the mine or a part of the mine be halted;

(b) the performance of any act or practice at the mine or a part of the mine be suspended or halted, and may place conditions on the performance of that act or practice;

(c) the employer must take the steps set out in the instruction, within the specified period, to rectify the occurrence, practice or condition; or
(d) all affected persons, other than those who are required to assist in taking steps referred to in paragraph (c), be moved to safety" (DMR, 2010a).

Due to the continuous high levels of accidents in mines, inspectors exercise these powers whenever non-compliances with safety regulations are observed. Many mines receive on a regular basis, what are known in the industry as *Section 54's*, and then efforts are made to rectify the hazardous situation as soon as possible in order to resume mining operations.

A current challenge within the MHSA is that, with such high levels of injuries and deaths (as reported in Figure 4.2 in Section 4.4), many inspectors are spending more time on reactive work, such as accident investigations and inquiries, than on proactive work like inspections and audits (PMHSA, 2008).

4.3.2. Mine Health and Safety Council

The Mine Health and Safety Council (MHSC) is a tripartite council that draws mining expertise from employers in the mining industry, mine employees through their unions, and state representatives, in order to advise the Minister of Mineral Resources on all occupational health and safety issues in the mining industry relating to legislation, research and promotion. The MHSC's responsibilities are governed by Sections 43 and 44 of the MHSA and include the promotion of health and safety in the mining industry, the management of research in relation to health and safety in the mining industry, and liaison with other bodies concerned with health and safety issues (MHSC, 2010).

Mines are required to pay a safety risk levy which is used for health and safety research under the auspices of the MHSC. The Safety in Mines Research Advisory Committee (SIMRAC) was established in terms of Section 29 of the Minerals Act (Act 50 of 1991), with the principal objective of advising the Mine Health and Safety Council on the determination of the safety risks on mines, and the need for research into such safety risks (DST, 2010).

4.3.3. Chamber of Mines

The Chamber of Mines of South Africa (CoM) was founded in 1889 as a voluntary membership, private-sector employer organisation. Members of the Chamber employ more than 400 000 workers, and some of the smaller mines are not members. The Chamber advocates major policy positions endorsed by the mining employers and represents these to government and other relevant policy-making and opinion-forming entities. The Chamber aims to be the recognised authoritative voice of mining in South Africa, whereby all significant stakeholders will regard the Chamber and its staff as the leading knowledge-based resource on all issues pertinent to mining (CoM, 2012a).

A current initiative of the Chamber is the Mining Industry Occupational Health and Safety Learning Hub, which encourages top health and safety performers to share their leading practices with other mines (Seggie, 2010). However, the Chamber identified that sharing best practices and learning from 'pockets of excellence' does not address the challenge of general adoption of the leading practices across the industry. Hence, they implemented an adoption system on a pilot basis and its members agreed to establish a Learning Hub to provide a more permanent system to support future adoption of leading safety practices. To date, four leading practice adoption teams have been established in the areas of Noise, Dust, Transport and Machinery, and Falls of Ground (Van der Woude, 2010). Furthermore, the Head of Safety at the Chamber, Mr. Sietse van der Woude, announced that the current system of production bonuses will be replaced by zero-harm operations bonuses to enhance safe operations (Van der Woude, 2012).

4.3.4. Association of Mine Managers of South Africa

The Association of Mine Managers of South Africa (AMMSA) represents employers in the South African mining sector and acts as a forum for discussion and evaluation of technical ideas and minerals policy issues. The vision of AMMSA is to promote and uphold the general advancement of the mining industry, with the safety and health of all mining employees under the members' management regarded as of paramount importance (AMMSA, 2012).

AMMSA co-signed a commitment to achieve the occupational health and safety targets and 10-year milestones agreed to at the 2003 Mine Health and Safety Summit.

- Gold Sector: By 2013, to achieve safety performance levels equivalent to current international benchmarks for underground metalliferous mines, at the least.
- Platinum, Coal and other sectors: By 2013, to achieve constant and continuous improvement equivalent to current international benchmarks, at the least (AMMSA, 2012).

To reach these milestones a reduction in fatality rates of 20% per year was required (CoM, 2007).

4.3.5. Mining Qualifications Authority

The Mining Qualifications Authority (MQA) is a statutory body established in terms of the MHSA and is a registered Sector Education and Training Authority for the Mining and Minerals Sector in terms of the Skills Development Act No 97 of 1998. Core functions of the MQA include the development and implementation of unit standards and qualifications for the Sector as well as overseeing the quality of standards, qualifications and learning provision.

The MQA entered into a Memorandum of Agreement with the Chamber of Mines in which it is agreed that the Chamber will coordinate the development of learning materials for the unit standard-based learnerships on behalf of the MQA. Training providers accredited by the MQA will do the actual development of the learning materials (MQA, 2014).

Application to Training

According to the MHSA, mines are legally obliged to compile various mandatory Codes of Practice (CoP). These CoPs are compiled in response to guidelines issued by the government department supervising mining activities, the Department of Mineral Resources. Failure to prepare and implement a CoP in compliance with a guideline is a breach of the MHSA. Consequently, all mines have significant numbers of CoP documents, also referred to as the mine standards. These standards usually make up most of the knowledge components that are presented in training courses at mines. In this regard, VR training can provide visual, interactive simulations of such content in order to improve the current training methods.

Tichon and Burgess-Limerick (2011) report that an integral aspect of improving mine safety has been an increased focus on ensuring that workers are competent to perform their duties and properly trained in the actions to take in dealing with hazards, or when unplanned events with adverse safety consequences occur. Failure to perceive a hazard is frequently identified as contributing to injuries and fatalities (Burgess-Limerick & Steiner, 2006; Kowalski-Trakofler & Barrett, 2003; Schofield, Hollands & Denby, 2001; Tichon & Burgess-Limerick, 2011), hence the VR prototype systems developed during this study focus on hazard recognition and rectification.

4.4. Mine safety statistics

During the past century, over 69 000 mine workers died and over a million have been seriously injured as a result of South African mining operations (Barry,1995:65; Krige,1995:9). To improve the conditions in South African mines, the Mine Health and Safety Inspectorate decided to collect detailed data on each accident in order to use the information for research purposes. The South African Mines Reportable Accidents Statistical System (SAMRASS) database was therefore established and has been active since 1988. In terms of the requirements of the MHSA, employers must report certain accidents and dangerous occurrences to the regional Principal Inspector of Mines, which is then captured into the SAMRASS database (DMR, 2012).

South Africa's gold mines are the world's deepest, with AngloGold's Savuka mine operating at a depth of 3.8 km. Mining at such depth can lead to earth tremors, which can result in rock falls. The worst local gold mine accident occurred in May 1995 at Anglo's Vaal Reefs mine when a runaway ore train severed a cage cable, causing 105 workers to plunge nearly 2 kilometres to their deaths (Sguazzin & Lourens, 2007). In 2007, 221 miners died in mine accidents in South Africa. The National Union of Mineworkers (NUM), the country's biggest labour union, called a one-day strike and led a protest march in downtown Johannesburg in December 2007 (Business Report, 2008).

From the earlier sections in this chapter, it is clear that within the South African mining industry regulatory environment, great emphasis is currently placed on creating a safe, healthy and productive work environment. Regardless of the regulatory framework, accident and injury statistics remain unacceptably high.

The next five graphs display data from the SAMRASS database (DMR, 2014a). Figure 4.2 shows the number of mining fatalities per year for the period 2003–2013, as was reported to the Department of Minerals and Energy (now the Department of Mineral Resources).

As indicated in Figure 4.2, there has been a steady decrease in fatalities over this period, except for 2007, but the mining industry still has to make significant safety advances in order to reach the zero-harm target. On average, over the 11 years, and after the mining industry committed in 2003 to strive towards zero harm, approximately 14 miners were killed in the South African mining industry every month.

Figure 4.2: Mining fatalities per year in the South African mining industry for the period 2003–2013.

Figure 4.3 shows the total accidents per month for the period January 2008 to June 2013, with the blue trend line showing the moving average over a twelve-month period. Similarly, Figure 4.4 shows the number of fatal accidents per month for the same period, with a yellow trend line indicating the moving average over a twelve-month period.

From Figure 4.3 it is clear that December 2009 was an extraordinary month with 95 accidents, while the lowest number of accidents (15) occurred in June 2013. The blue trend line did not deviate much between middle 2011 and middle 2013, and reached the same level in June 2013 as in April 2011. The yellow trend line in Figure 4.4 reached its lowest point in June 2011, but only reached similar levels again in mid 2013. From the three figures (Figures 4.2 to 4.4) it can be seen that total accidents and fatalities have decreased over the period 2008–2013. However, the number of total accidents and the number of fatal accidents have not decreased significantly over the last two years (2011–2013). Despite this, the number of fatalities dropped from 123 in 2011 to 93 in 2013. It is, however, a cause for concern that the number of accidents in the recent past remained constant. It is clear that additional interventions are required to ensure a further, more drastic, reduction in accidents.

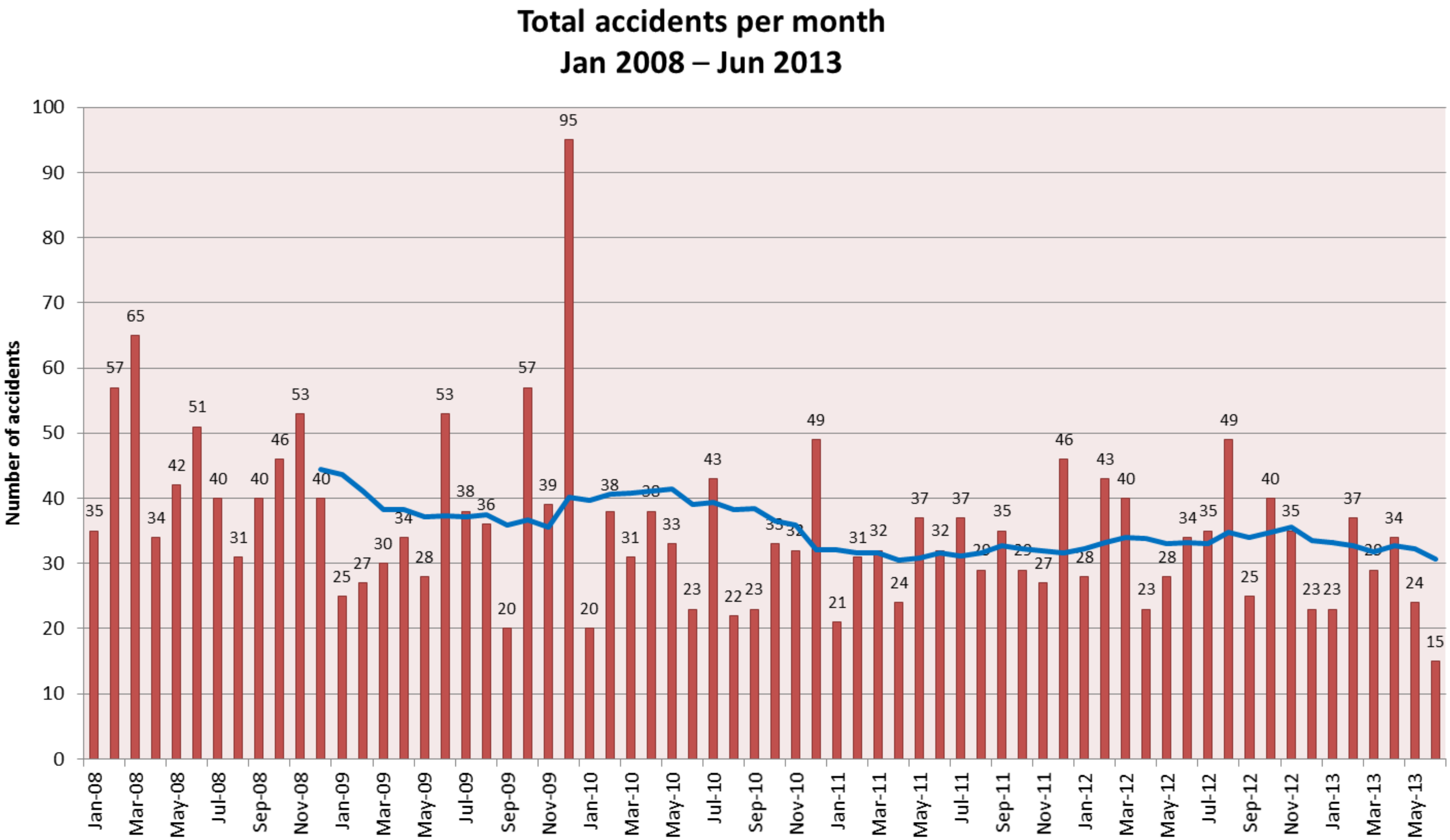


Figure 4.3: Total accidents per month in the South African mining industry for the period January 2008 to June 2013.

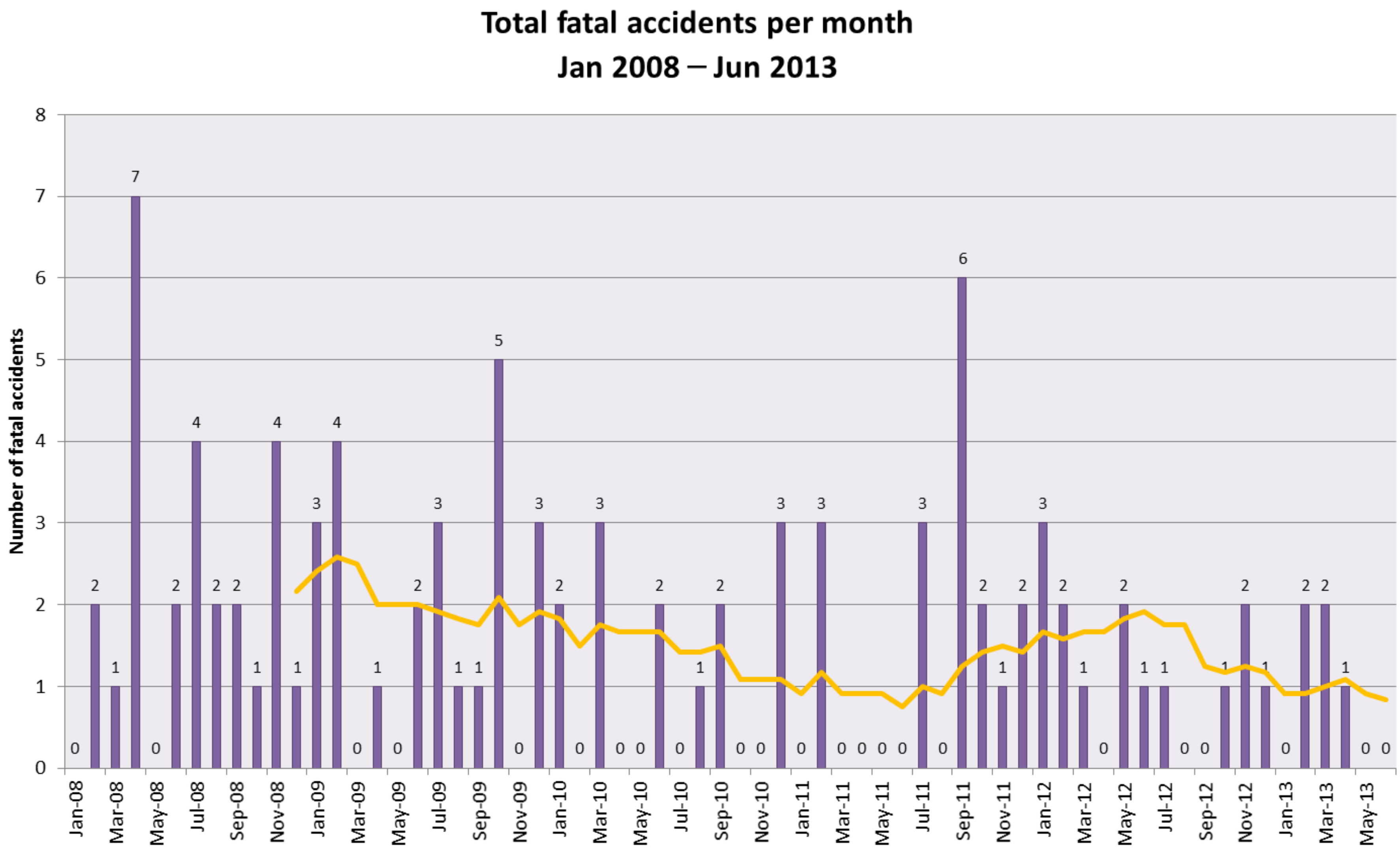


Figure 4.4: Total fatal accidents per month in the South African mining industry for the period January 2008 to June 2013.

Figure 4.5 shows the percentage of fatalities per commodity type and Figure 4.6 the percentage of fatalities per fatality category (the cause of death) for the period 2006 to 2013. It is clear from Figure 4.5 that the gold mines have the highest fatality rate (51%) compared with other commodities mined, and that 80% of the fatalities occur in hard rock mining (gold and platinum).

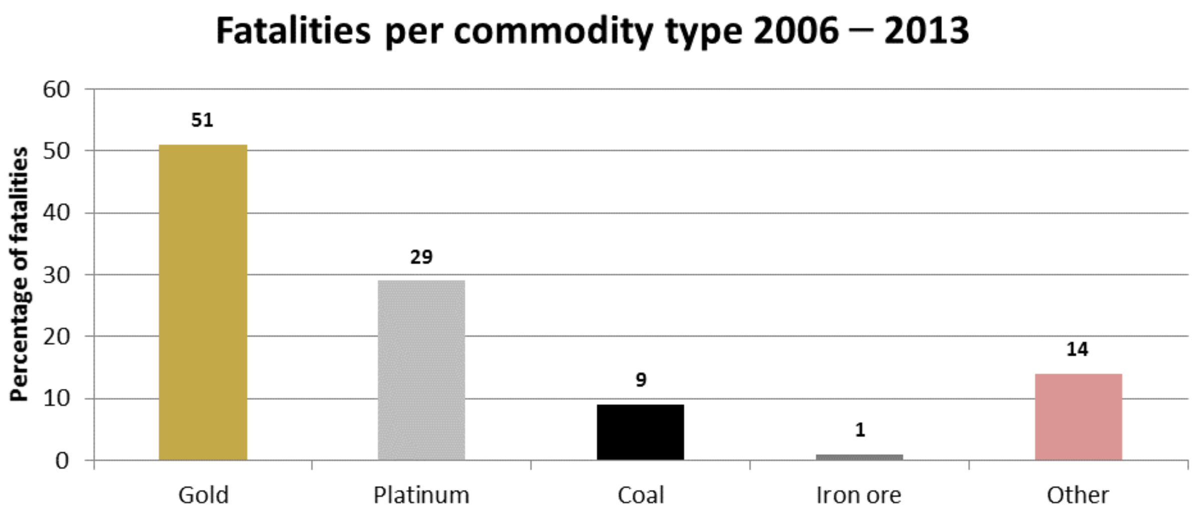


Figure 4.5: Percentage distribution of fatalities per commodity type for the period 2006 – 2013.

In Figure 4.6 it is shown that, for the period under consideration, 37% of the fatalities were related to rock falls (falls of ground are referred to as FOGs), which remain one of the major sources of potential underground danger to mine workers. The major contributor to fatalities (FOGs) is closely followed by trackless mechanised mining machinery (called TM3, 32%), while accidents relating to general hazards contributed 21% to the overall fatality statistics.

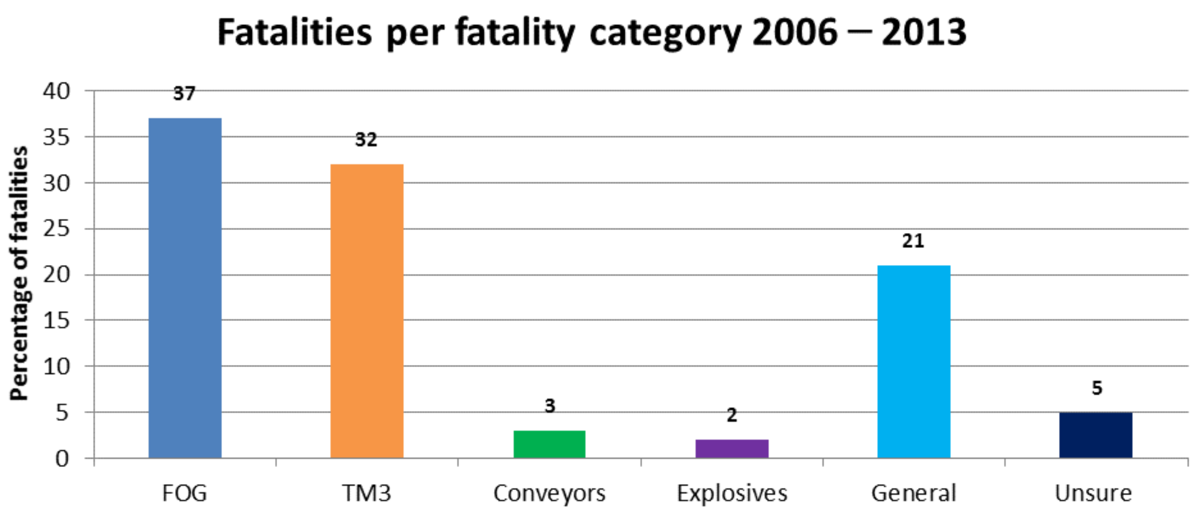


Figure 4.6: The percentage distribution of fatalities per fatality category for the period 2006 – 2013.

Figure 4.7 indicates the mining industry performance versus the 2013 milestones, as explained in Section 4.3.4. The targets were set in terms of fatality frequency rate (FFR), measured per million hours worked. It is clear that the actual performance came close to the targets in 2004 and 2005, but thereafter the gap widened and the milestones were not reached.

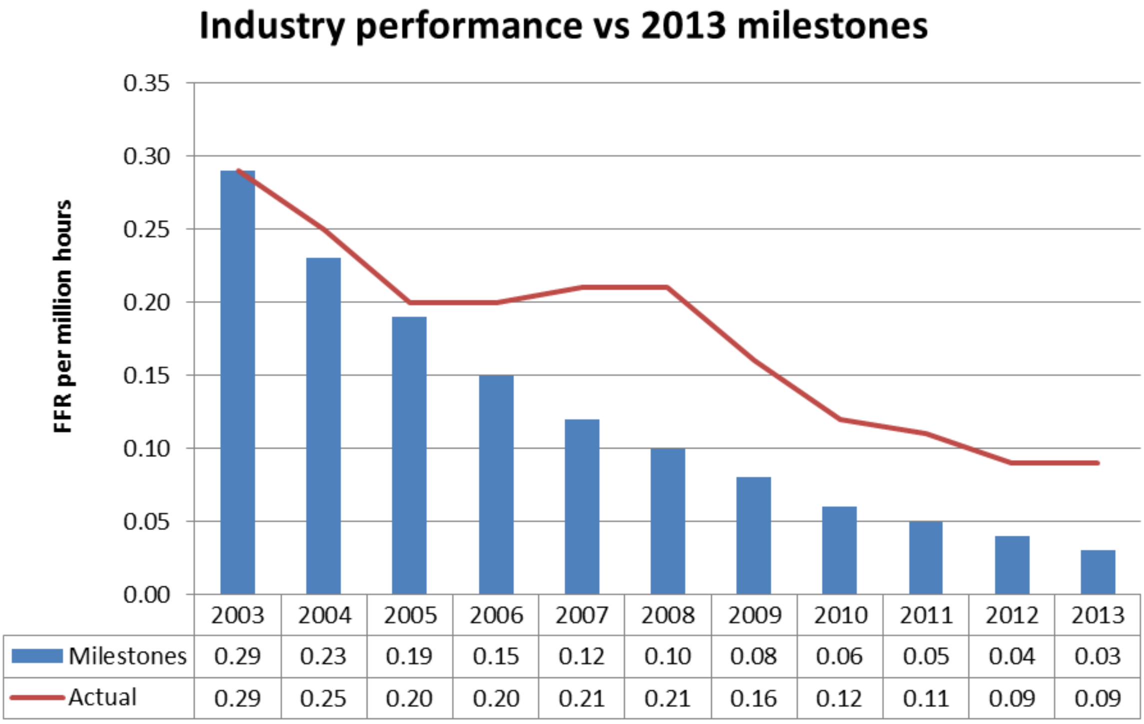
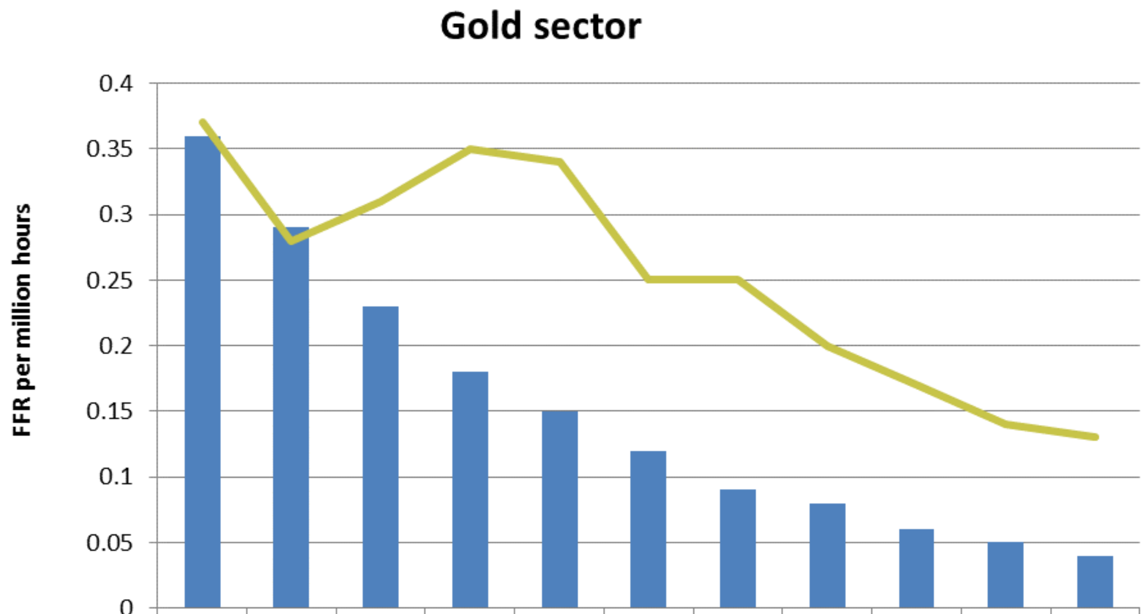


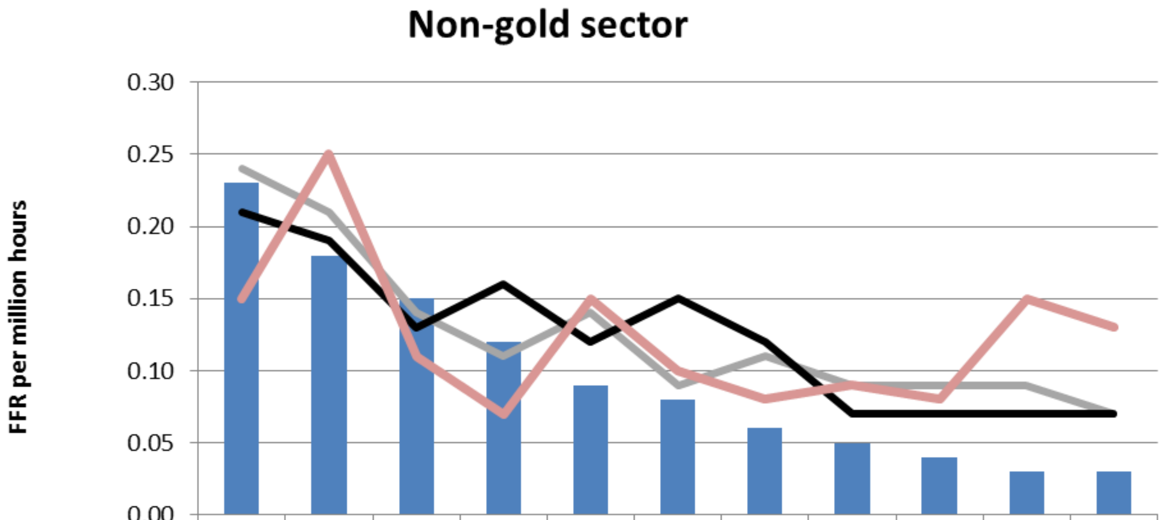
Figure 4.7: Mining industry performance versus 2013 milestones (Briggs, 2014).

Since different milestone targets were set for the gold mines, Figures 4.8 and 4.9 distinguish between the gold sector and the other mining sectors by showing the performance of the gold sector versus the 2013 milestones in Figure 4.8 and the performance of the non-gold sector in Figure 4.9. Figure 4.8 shows that, even though the gold sector reached its milestone target in 2004, the period 2005–2007 had notable increases in fatalities. After a stabilisation period (2008–2009), the fatalities did decrease in following years, but still ended up being more than three times higher than the 2013 target (0.13 instead of 0.04).



	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Milestones	0.36	0.29	0.23	0.18	0.15	0.12	0.09	0.08	0.06	0.05	0.04
Actual Gold	0.37	0.28	0.31	0.35	0.34	0.25	0.25	0.20	0.17	0.14	0.13

Figure 4.8: Gold sector performance versus 2013 milestones (Briggs, 2014).



	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Milestones	0.23	0.18	0.15	0.12	0.09	0.08	0.06	0.05	0.04	0.03	0.03
Actual Platinum	0.24	0.21	0.14	0.11	0.14	0.09	0.11	0.09	0.09	0.09	0.07
Actual Coal	0.21	0.19	0.13	0.16	0.12	0.15	0.12	0.07	0.07	0.07	0.07
Actual Other	0.15	0.25	0.11	0.07	0.15	0.10	0.08	0.09	0.08	0.15	0.13

Figure 4.9: Non-gold sector performance versus 2013 milestones (Briggs, 2014).

Figure 4.9 indicates the performances of the platinum and coal sectors (in silver and black respectively) and group the performances of the other mines together (shown in pink). This graph shows that the platinum and coal sectors had similar decreases in their fatality rates, but the other mines had two disastrous years in 2012 and 2013. None of the categories came close to the 2013 targets, with platinum and coal still reporting more than double the number of fatalities than the target (0.07 instead of 0.03), while the other mines ended 2013 with a fatality rate more than four times higher than the target (0.13 instead of 0.03).

In terms of overall performance, the biggest gap between actual performance and the milestones is on 'Other' commodities, which have only reduced their fatality rate by 13% over the last decade. The platinum sector, which is the application domain of the present research, has reduced its fatality rate by 71%, followed by coal (66%) and gold (64%). This indicates a tenacious and supportive environment as context for this research.

A study by the Chamber of Mines in 2004 showed that the safety performance of the South African mining industry is more than 50% worse than Australia, Canada and the USA. An annual improvement of 20% per year would have been required to reach those international levels by 2013 (Seccombe, 2007). From the sector performances shown in Figures 4.8 and 4.9, only the coal sector now has a fatality frequency rate comparable to the USA (Briggs, 2014).

During 2009 it was reported that, even though the sector committed itself to achieve the 2013 targets, if the slow rate of safety improvements of the previous decade was maintained, the industry would miss its 2013 targets by 24 years (Brown, 2009). Even if fatalities are reduced to the international benchmark of one death for every 33 million hours worked, fatalities would be far from eliminated as this would still mean that 34 people would die in mining per year.

Application to Training

Due to the complexity of South African deep-level mining, knowledge (what to do) and skills (how to do it) are essential drivers of good health and safety performance. "There is no doubt that improvement of health and safety knowledge will contribute greatly to the reduction of injuries and deaths at mines" (PMHSA, 2008:35). The rate of fatalities and the opinions of experts all point in a single direction:

Improvements in safety training are required to reduce the human error factor.

The general low literacy levels of employees in the mining sector compound the problem of safety and health training. Table 4.3 shows the educational profile of the mining industry, comparing the 2007 data to 2013. As Table 4.3 indicates, in 2007 almost a quarter of mine employees had no basic education and, on average, 80% had below grade twelve education. The MQA is responsible for facilitating skills development in the mining industry. One of their challenges is to eliminate illiteracy in the mining sector, which is tackled through Adult Basic Education and Training (ABET). The 2013 data presented in Table 4.3 shows considerable improvement over the 2007 figures, with less than 13% of the work force having no school education, compared to 23% in 2007. Employees with matric (Grade 12) qualifications increased from 15% in 2007 to 26.1% in 2013, and the number of employees with post-matric qualifications more than doubled from 6.6% to 13.7%.

A further challenge in the workforce lies within the realm of communication. Effective communication is essential for the safe and healthy operation of the industry. The workforce speaks a range of languages and the mining lingua franca, called *Fanakalo*, is inadequate to convey the nature and extent of hazards. Lack of literacy renders written communication impossible with a large portion of the workforce. In order to undertake meaningful theoretical and textual training courses, a major upgrade of educational levels would first need to be achieved. This would take many years and is a further factor supporting the need to use visualisation and simulation in safety training.

Table 4.3: Educational profile of the mining industry (PMHSA, 2008; MQA, 2014).

Educational Level	**Percentage of employees (2007)**	**Percentage of employees (2013)**
No schooling / pre-ABET	22.8	12.7
Grade 3 / ABET 1	6.3	4.6
Grade 5 / ABET 2	9.9	6.9
Grade 7 / ABET 3	12.6	8.7
Grade 9 / ABET 4	13.6	7.5
Grade 10	7.1	9.3
Grade 11	6.1	10.5
Grade 12	15.0	26.1
Post school qualifications	6.6	13.7
TOTAL	100	100

Matthee, Henneke and Johnson (2014) highlight the factors that contribute to the resistance to, and/or adoption of, e-learning as perceived by e-learning managers and practitioners in the mining industry. Some of their findings indicate that:

- quality content is a necessary condition for successful adoption,

- expectations of different stakeholders involved in e-learning should be properly communicated, and
- a focus on people development, rather than only on legal compliance, may lead to lower resistance.

This section has highlighted the need for enhanced training in terms of:

- safety training,
- basic education, and
- communication skills,

and the role that visualisation and simulation can play.

4.5. Competency certificates

Government certificates of competency are required for appointment at management and supervisory levels in the mining sector. As Webber-Youngman and Van Wyk (2009) point out, the knowledge level of trainees coming through the system to fulfil the need in terms of addressing health and safety challenges that are currently facing the mining community, is alarming.

4.5.1. The Mine Manager's Certificate of Competency

The examination for the Mine Manager's Certificate of Competency can be written in May and October every year. The examination consists of three parts with a total of seven papers: Part A covers the topics of engineering, geology and surveying; Part B has three mining-related examination papers; and Part C is a paper on legal aspects pertaining to the mining industry. Table 4.4 shows the subject average results of the examinations that were written between May 2009 and May 2012 (DMR, 2014c).

From Table 4.4 it is clear that a major challenge exists in terms of the quality and level of knowledge of the students that sit for this examination. The only average percentage that reached a passing grade is indicated in green font and occurred in the Mining 3 paper of May 2012. These poor averages also raise the issue of the competency level of the candidates who attempt these examinations, as several 'zero' percentages have been recorded in all the subjects mentioned above. The overall average mark obtained for the Surveying paper was only 16%, with Engineering having the second-worst average result of 26%. Both these subjects require a reasonable ability to apply mathematical skills and the suitability of candidates is an area of concern. The overall average for all the students in all seven subjects is only 32%.

Table 4.4: Mine Manager's Certificate of Competency subject average results for the period May 2009 to May 2012.

	PART A			PART B			PART C	Average
	Engineering	Geology	Survey	Mining 1	Mining 2	Mining 3	Legal	
May 09	12%	38%	18%	27%	18%	35%	30%	**25%**
Oct 09	32%	40%	13%	32%	24%	22%	43%	**29%**
May 10	23%	47%	23%	30%	42%	40%	45%	**36%**
Oct 10	36%	44%	8%	32%	35%	43%	44%	**35%**
May 11	27%	20%	14%	37%	26%	39%	43%	**29%**
Oct 11	38%	30%	13%	34%	21%	40%	47%	**32%**
May 12	16%	32%	25%	41%	37%	55%	38%	**35%**
Average	**26%**	**36%**	**16%**	**33%**	**29%**	**39%**	**41%**	**32%**

Figure 4.10 shows the number of students who wrote each exam (in blue), as well as the number of students that passed each exam (in red). All the subjects show poor pass rates, especially with only 28 of 390 students passing Surveying, which is a 7% pass rate. In terms of the Legal subject, only about a third of the candidates were successful, which highlights a serious area of concern in terms of the mineworkers' comprehension of legal aspects pertaining to mining.

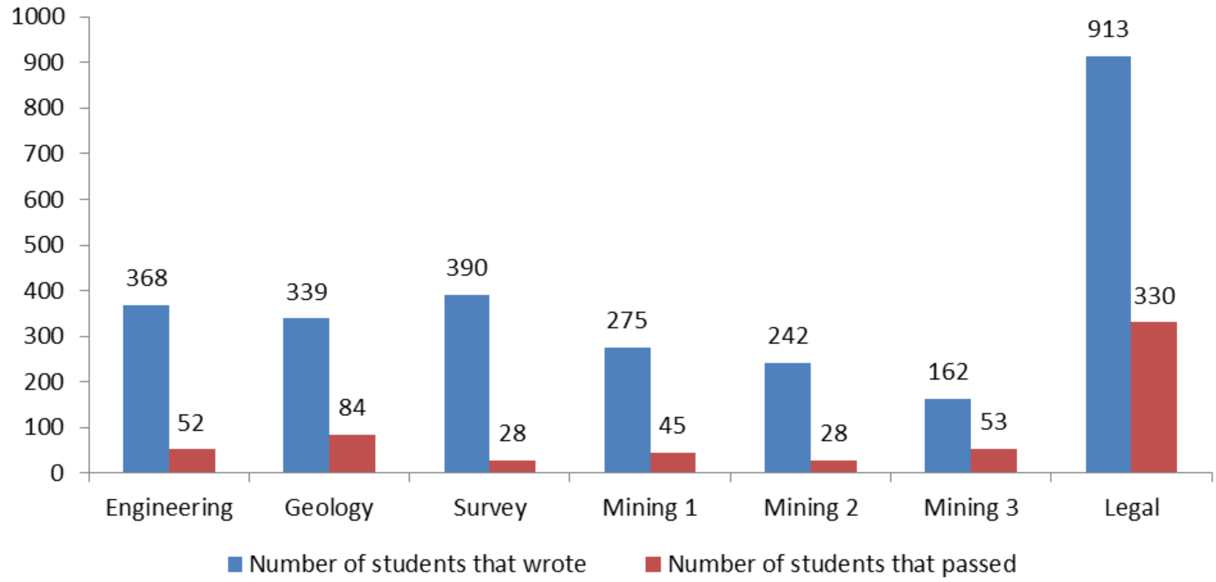


Figure 4.10: Mine Manager's Certificate of Competency individual results for the period May 2009 to May 2012.

These poor results demonstrate the need to introduce innovative education strategies to raise pass rates to acceptable levels, yet without compromising standards (Webber-Youngman & Van Wyk, 2009).

4.5.2. The Certificate in Rock Engineering

Many of the certified rock engineers who qualified through the Chamber of Mines over the last few decades left South Africa to take up employment abroad. These losses will have to be replaced with up-and-coming young rock engineers, which is a major challenge facing the mining industry in South Africa. Registration for the Chamber of Mines Certificate in Rock Engineering also shows a serious downward trend, causing a shortage in skilled rock engineers (Webber-Youngman & Van Wyk, 2010).

With rock falls being the main contributor to mining fatalities, training in strata control, that is, correctly supporting the hanging wall after excavation, should receive high priority. In a SIMRAC study by the Council of Scientific and Industrial Research (CSIR) Division of Mining Technology, Hagan (2004:33) concludes that "it is clear from several in-depth assessments of accident records that a better understanding of very basic strata control principles and practices, by production personnel, will improve worker safety". It is therefore cause for concern that Venter (2007:4), in a research report conducted at Harmony Gold, states that "most training instructors were found to lack the necessary understanding and basic knowledge to effectively train in rock strata. It was also found that some instructors are using their own theories to explain some rock engineering related theories, which were totally incorrect".

The issues identified in this subsection pinpoint knowledge problems at the level of professionals. This is over and above the training requirements of mine workers, which is a primary focus of this research.

4.6. The Presidential Mine Health and Safety Audit

Due to the number of serious accidents during 2007, the then State President requested the Department of Mineral Resources (the then Department of Minerals and Energy) to conduct a full audit of all mines. The main aim of the audit was to determine whether the mines meet health and safety standards as prescribed by the Mine Health and Safety Act (PMHSA, 2008).

The audit was prompted after an accident at Harmony Gold's Elandsrand mine on 4 October 2007, when 3200 workers were trapped underground for 30 hours before being rescued. This incident followed soon after four workers died at AngloGold Ashanti's Mponeng mine due to a rock fall 3.3 km below the surface (Seccombe, 2007). The audit process involved scrutinising the health and safety management systems of the mines and conducting inspections to verify their effectiveness. Between December 2007 and May 2008, the Mine Health and Safety Inspectorate (MHSI) audited 355 mines. The consolidated audit results of all the mines revealed an overall compliance rate of only 66%, which is inadequate.

The audit confirmed the findings of the Leon Commission prior to the MHSA, namely, that "there is a pervasive culture of non-compliance to legislative requirements. Inquiry after inquiry makes findings to the effect that risk assessments are not conducted, training is not done, early morning examinations are not done, equipment not maintained and the list goes on and on" (PMHSA, 2008:51).

Although the MHSA makes provision for administrative fines for non-compliance, this failed to serve as a deterrent, because the maximum fine of R200 000 (approximately $26 000) at the time of the audit, was relatively low and mines continuously appealed the fines. It can take the MHSI up to 18 months to process a single fine. The MHSA also makes provision for prosecution of managers in cases of negligence, but no prosecution has yet taken place, even though several charges have been laid. As a result of the audit report, the Mine Health and Safety Amendment Act of 2008 increased the fines for non-compliance to health and safety laws to R1 million (approximately $130 000). However, such a fine could still be easily paid by major mining companies. In comparison, British Petroleum (BP) was fined $87 million (approximately R672 million) in the USA after an explosion at its Texas refinery killed 15 people (Brown, 2010). Furthermore, negligent officials can now face up to five years in prison (Badenhorst, 2011).

The Chamber of Mines reacted to the Amendment Act describing it as counterproductive. In their view the punitive measures will place blame after an accident or fatality, but are unlikely to succeed in preventing accidents and fatalities. The correct balance between punitive and preventative measures should be found (Pringle, 2009).

The presidential audit report particularly mentions the gold and platinum mines as serious concerns. This explains the tough stance currently taken by the inspectors to stop operations with Section 54 notices at workplaces where there is non-compliance, as explained in Section 4.3.1. Even though this is costly and disruptive to mines, inspectors use this regulatory approach as mine managers need to grasp implicitly that the current high rates of injuries and deaths are unacceptable. As examples, Gold Fields lost R2.3 billion during 2008 due to safety stoppages and improved safety measures. Moreover, it was estimated that Impala Platinum's accident in July 2008, when 9 people died, cost the mine 150 000 ounces of platinum production worth R1.5 billion over the five years following the accident, due to stoppages and remedial action (Brown, 2009). During 2013 a total of 1074 Section 54's were issued to mines in terms of the MHSA to protect health and safety of mineworkers and to address unsafe or hazardous conditions, practices or acts (DMR, 2014b).

Application to Training

The findings in the previous paragraphs clearly indicate that there is a lack of formal training on hazard identification and risk assessment. At many mines there is a lack of infrastructure to ensure effective training provision. Accidents and fatalities have impacts on productivity and financial loss that extend far beyond the initial tragedies. Means must be sought to halt this process. The presidential audit report recommends that health and safety training should be prioritised to increase the level of safe behaviour at workplaces.

This study addresses the above-mentioned lack of training on hazard identification by proposing the use of virtual reality to improve training in hazard recognition and remedial actions. Two VR training prototypes were developed for this purpose. These prototypes simulate the underground working areas, incorporating potential hazards that mine workers need to identify and indicating possible actions that might be followed in response to each hazard. The development of these prototypes is discussed in Chapters Six and Seven.

Furthermore, with the aim of developing improved training interventions, it was decided to undertake a case study to determine more details on causes of accidents. Details of the case study are also presented in Chapter Seven.

4.7. Conclusion

This chapter briefly introduced the South African mining environment in Section 4.1. Section 4.2 presented an overview of the safety legislation applicable to the mining industry, and Section 4.3 discussed the major stakeholders involved in this industry.

The Mine Health and Safety Act (MHSA, 1996) recognises the need to improve the level and effectiveness of training given to underground workers. The legislation requires that all mine workers receive appropriate and effective training to enable them to operate safely in their environments.

In Section 4.4 mine safety statistics were presented. The significance of these statistics is that there has been a steady decrease in fatalities through sustained efforts by mining companies. Regardless of the regulatory framework and stringent measures to deal with non-compliance, the numbers of fatalities and serious accidents in the mining industry are still unacceptably high, and more emphasis should be placed on strategies to achieve the zero-harm goal embraced by all South African mines. Lack of knowledge (through a lack of education and training) plays a definite role in increasing the fatality rate (Webber-Youngman & Van Wyk, 2013). Various approaches should therefore be employed to make workers aware of the consequences of unsafe acts, and here VR technology has an important role to play in reinforcing awareness of these consequences.

Section 4.5 discussed the results of industry competency certificate examinations. Falls of ground still remain the main reason for fatalities on South African mines. Moreover, the exodus of rock engineering practitioners from South Africa over the last 17 years is an area of concern in that there are insufficient competent rock engineers to deal with the challenges associated with rock engineering (with specific reference to falls of ground).

The presidential audit highlighted the need for improved training. Current training methods rely on repetitive classroom style learning, followed by on-the-job training. Some mines have a physical mock-up of an underground workplace as a training facility. Improved methods are required to provide effective safety training. This research will show that the application of VR training systems is an innovative approach to safety training in the mining industry that holds potential for reducing fatalities and injuries.

At a safety *indaba* held in 2007, Minerals and Energy Minister, Buyelwa Sonjica, confirmed that rockfalls, rockbursts and seismicity were the biggest contributors to fatalities and injury. The Mining Inspectorate emphasised that a poor safety culture was a concern, as some workers failed to recognise hazardous conditions. In studying the mining safety record in 2006, the Chamber of Mines found that human factors contributed to most fatalities. Although shortages in scarce skills such as rock engineering was a contributing factor, the main concerns indicated were the poor safety culture and the fact that human factors require considerable attention (Ndaba, 2008).

Chapters Six and Seven detail the design, development and evaluation of VR systems to improve hazard recognition and strata control training, thereby contributing towards addressing the problem of accidents caused by human factors. The next chapter, Chapter Five, presents the research design of this study.

Chapter Five
Research Design and Methodology

5.1. Introduction

The previous three chapters reviewed literature relevant to this study. This chapter, by contrast, looks forward by explaining the foundations and processes of this research. A research study of this nature necessitates careful planning. This chapter defines the underlying research paradigm of this study, the research and data collection methods used and the research design of the study. Furthermore, a heuristic framework is proposed for the evaluation of desktop VR training systems.

Section 5.2 revisits the research questions given in Chapter One, indicating where each research question is addressed in the thesis. The primary research paradigm of this study is *design research,* which is currently a maturing research methodology within a number of disciplines. In the realm of educational technology it is called *design-based research.* Design research is also increasingly used for studies in information systems research, where it is termed *design science research*. Section 5.3 introduces design science and design research. Design science research is overviewed in Section 5.4 while Section 5.5 explains the origin, philosophical underpinnings, characteristics and principles of design-based research. The research design of this study is discussed in detail in Section 5.6. The researcher's synthesised model of design-based research is presented, along with an elaboration of how four cycles of design-based research were applied. This led to the implementation of a real-world solution in the form of two VR safety training systems, and the generation of theory in the form of an evaluation framework, which implicitly could also serve as a set of design principles. A set of interrelated diagrams depicts the flow of the research processes. This research was conducted at a large platinum mine in Rustenburg, South Africa, which will hereafter be referred to as the Mine.

Section 5.7 highlights the research methods used and indicates where each method was applied in the study. Section 5.8 presents the evaluation framework proposed by this study, which comprises four categories of evaluation criteria. This heuristic evaluation framework is applicable to the evaluation of desktop virtual reality training systems and is referred to as the desktop VR evaluation framework (*DEVREF*). The *DEVREF* Framework was applied in the evaluation of two prototypes, details of which are discussed in Chapter Eight. Finally, Section 5.9 considers the aspects of validity, reliability and triangulation, while Section 5.10 highlights ethical considerations of the study. Figure 5.1 graphically indicates the layout of this chapter.

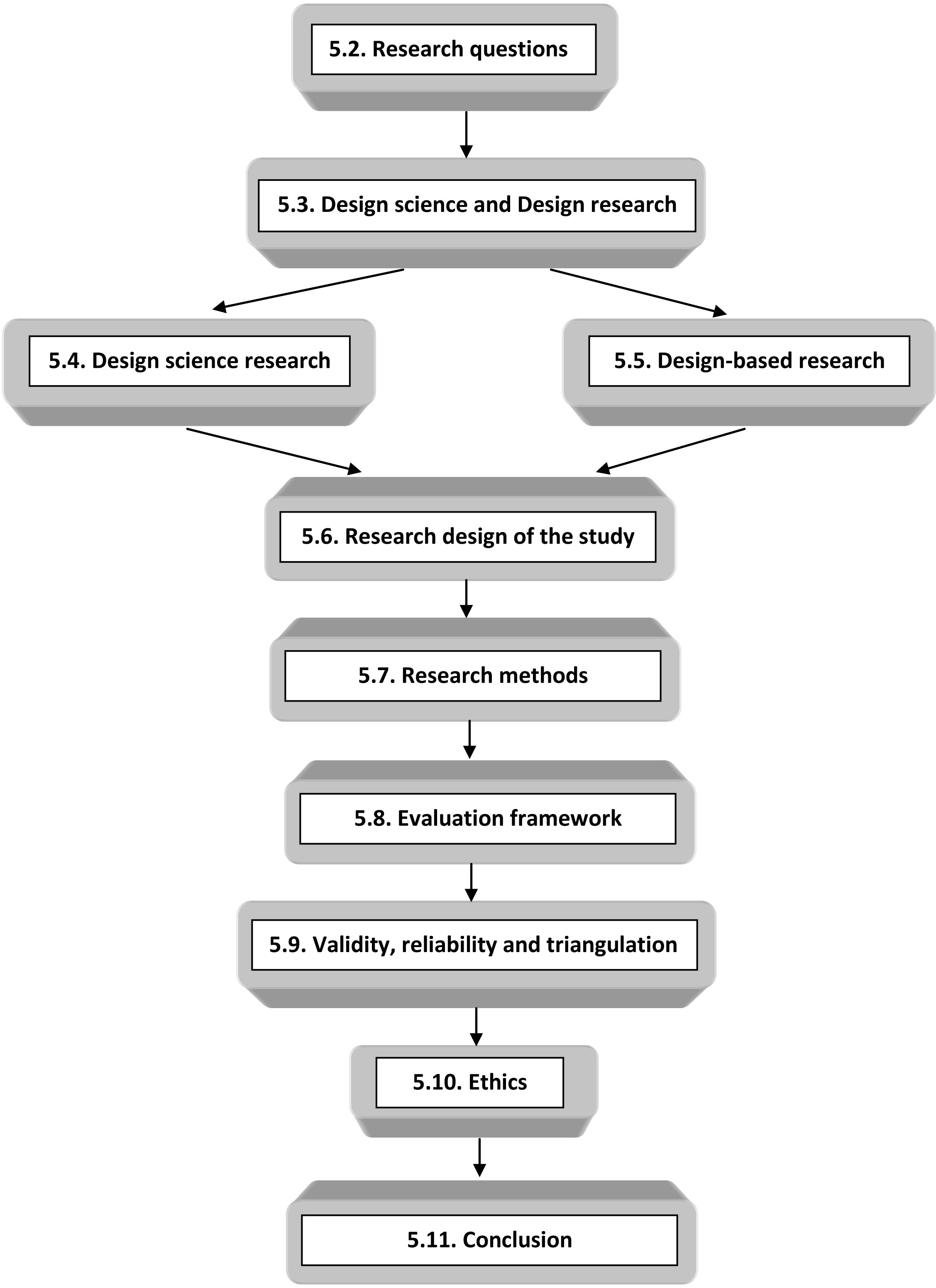


Figure 5.1: Layout of Chapter Five.

5.2. Research questions in the context of the study

The primary aim of this research was to generate an evaluation framework to evaluate interactive desktop VR training systems. The secondary aim was to propose, design and develop novel e-training interventions on the topic of safety training for mineworkers operating in the underground mining environment. The training would be delivered by desktop VR technology. To achieve the stated aims, a main research question and six subquestions were formulated.

As stated in Section 1.4 in Chapter One, the Main Research Question addressed by this study, is:

What is an appropriate and effective framework for evaluating virtual reality training systems in the mining industry?

To realise the stated aims, a number of research subquestions need to be addressed. Table 5.1 revisits the research subquestions and indicates in which section(s) in the study each subquestion was addressed.

Table 5.1. Research subquestions with corresponding locations in thesis.

#	Research subquestion	Location in thesis	
RQ1	What is the suitability and potential of virtual reality technology for training applications in the domain of mine safety training?	Chapter 2	Sections 2.4 and 2.5
RQ2	Which research paradigm is appropriate for the intended research?	Chapter 5	Sections 5.5 and 5.6
RQ3	What are the contextual requirements for virtual reality training systems for the mining industry?	Chapter 6	Section 6.5
		Chapter 7	Section 7.2
RQ4	What is an appropriate design lifecycle model for interactive desktop virtual reality training systems?	Chapter 7	Section 7.3
RQ5	What structure, categories and criteria should be incorporated in an evaluation framework for virtual reality training systems in the mining industry?	Chapter 5	Section 5.8
		Chapter 9	Section 9.4
RQ6	How appropriate and effective is the proposed framework?	Chapter 9	Sections 9.2 and 9.3

5.3. Design science and Design research

Design research originated from the work of Simon (1981), who distinguished between the *natural sciences*, such as anatomy, astronomy, chemistry and physics, and the sciences of the artificial, or *design sciences,* such as engineering, product design, information technology and instruction. In the natural sciences, *descriptive* laws represent natural phenomena, while theories and formulae explain how they occur. The design sciences, by contrast, relate to man-made phenomena, where theories and models outline goals to be achieved and procedures to accomplish them, which are set out by *prescriptive* laws. Design science is characterised by the construction and evaluation of innovative artefacts and interventions in authentic settings.

Applied design science led to design research, which is called *design science research* (DSR) in the discipline of information systems (Hevner, March, Park & Ram, 2004) and *design-based research* (DBR) in the fields of educational technology and e-learning (Barab and Squire, 2004; Wang and Hannafin, 2005). These two forms of design research are discussed in the next two sections.

5.4. Design science research

In the discussion of information systems, design science research is used to address the creation and evaluation of IT artefacts specifically intended to solve real-world problems. The artefact should be effectively described to enable application and implementation in an appropriate domain (Hevner *et al*., 2004). The main focus is to produce and evaluate new and improved artefacts as solutions to complex problems and to generate new knowledge for the body of scientific evidence (Adikari, McDonald & Campbell, 2011).

In pioneering work, March and Smith (1995) proposed applying design research in the disciplines of information systems and information technology under the name design science research. They define the IT artefacts produced by DSR as *constructs*, *models*, *methods* and *instantiations*. Constructs are the basic concepts of a domain. When multiple constructs are combined to indicate relationships, the resulting artefacts are called models. Steps and processes that perform goal-directed activities are called methods. An instantiation involves actual implementation of a working artefact to demonstrate its feasibility and effectiveness. Rossi and Sein (2003) proposed a fifth output of design science research, namely *better theories.*

DSR can lead to theory building as the artefact exposes relationships between its elements which increase understanding of the elements and can result in elaboration of the previously theorised relationships. However, there is no general agreement on whether theory is indeed a required output. Gregor and Hevner (2013) suggest that partial or incomplete attempts at generating new theory via the development of artefacts are also acceptable, as this can indeed contribute to evolving theory. They also state that "evaluation of the artefact with quantitative measures of effectiveness provides the empirical evidence for theory development" (Gregor & Hevner, 2013:A1).

Markus, Majchrzak and Gasser (2002) advocated DSR within the context of information system design to prescribe effective development practices (methods) and a system solution (instantiation) for a particular class of user requirements (models). Design science research can be differentiated from pure design or system building due to the contribution design science makes to the knowledge base of foundations and methodologies. In the design of information systems, existing knowledge is applied to organisational problems, but design science research focuses on "addressing important unsolved problems in unique or innovative ways or solved problems in more effective or efficient ways" (Hevner *et al.*, 2004:81). Hevner *et al.* (2004:83) proposed seven guidelines based on the DSR principle that "knowledge and understanding of a design problem and its solution are acquired in the building and application of an artefact".

1. Design: An innovative, viable artefact must be designed and produced (construct, model, method or instantiation).
2. Relevance: A technology-based solution must be developed that is relevant to a specific authentic organisational problem.
3. Evaluation: Appropriate evaluation methods should be used to evaluate the quality, utility and efficacy of the design artefact.
4. Research contributions: Contributions should be clear, verifiable and innovative.
5. Research rigour: Rigorous methods should be applied in both construction and evaluation of the artefact.
6. Design as a search process: DSR is inherently iterative and design is an iterative search process to discover an effective solution. Reaching desired ends should be achieved through context-dependent abstractions and representations, while still satisfying laws in the problem space. This will invariably involve innovation and creativity.
7. Communication of results: Construction and evaluation details should be effectively presented to technology-oriented audiences, while management-oriented audiences will be concerned with the details relating to the artefact's impact, novelty and effectiveness.

Venable (2010) evaluated the rigour and relevance of design science research and found that the above-mentioned guidelines of Hevner were largely endorsed by researchers applying DSR, but caution was raised that these guidelines should be applied less mechanistically and that further improvements to criteria, standards and guidelines were required. Gonzalez and Sol (2012) stressed the importance of artefact evaluation and validation of the research process as elements that contribute to increased rigour, clarity and structure in design science research.

5.5. Research paradigm of this study: Design-based research

The next two sections address Research Subquestion 2 of this study:

RQ2	Which research paradigm is appropriate for the intended research?

Design research is increasingly used for studies within the context of educational technology, especially for studies on the development of e-learning and e-training (De Villiers, 2012). The term used to describe design research in this context is design-based research. Design research, in particular design-based research, was selected as the underlying research paradigm of this study because of its cyclic nature of design, evaluation and redesign, and its mandatory production of both theory and actual solutions in real-life contexts. An alternative paradigm could have been action research (De Villiers, 2005b; McNiff, 2013; Noffke & Somekh, 2009), since it is also iterative and can apply to inventions, interventions and products. However, DBR was deemed the most appropriate choice for this research due to its focus on:

(i) solving complex problems,
(ii) producing authentic artefacts, and
(iii) generating dual outcomes.

A broad definition of e-learning applications includes forms and methodologies such as tutorials, multimedia productions, simulations, educational games, interactive learning/practice environments, immersive virtual reality technology, educational software, Web-based learning applications, and learning management systems (Alessi & Trollip, 2001; De Villiers, 2005a). The design-based educational technology research approach can be applied to the delivery, content and architecture of such systems (De Villiers & Harpur, 2013). Wang and Hannafin (2005) state that design-based research has an important role to play in the development of technology-enhanced learning environments, which incorporate teaching for the acquisition of skills and knowledge, a variety of tools, and technological resources.

5.5.1. Origin of design-based research

Design-based research (DBR) has its roots in the design sciences such as engineering and product design, where iterative and context-based processes are followed to create usable products (Zaritsky, Kelly, Flowers, Rogers & O'Neill, 2003). DBR evolved as a research methodology based on the initial work of Brown (1992) and Collins (1992). Brown (1992) described using *design experiments* to bridge studies of learning with studies of instructional interventions in complex and changing environments. Collins (1992) proposed a more systematic methodology for conducting such design experiments that would assist in developing design theory to guide implementation of innovations. Van den Akker (1999) and Reeves (2000) used the terminology *development research*, but the Design-based Research Collective (2003) chose to use the term *design-based research* to avoid possible confusion with experimental design or experimentation with methods of teaching. In 2004, Barab and Squire (2004:12) stated that although design-based research as a term has grown in popularity and significance, "we are still at our infancy in terms of having agreement on what constitutes design-based research, why it is important, and methods for carrying it out". In their seminal paper they put forward particular assertions grounded in actual examples of their own and their colleagues' work. In another classic DBR paper, Wang and Hannafin (2005) defined characteristics of DBR and proposed principles for its application.

An alternative, more recent term for DBR used by some researchers is 'Educational design research' (Plomp, 2007; Teräs & Herrington, 2014; Van den Akker, Gravemeijer, McKenney, & Nieveen, 2006), but in this study the term *design-based research* will be used throughout.

5.5.2. Philosophical grounding of design-based research

Vaishnavi and Kuechler (2009) state that in any intellectual endeavour there are assumptions about the nature of reality (ontology), the nature of knowledge (epistemology) and values (axiology), but these assumptions are mostly implicit, even for researchers. They compared design research (design research as such, not necessarily DBR) to the positivist and interpretive approaches to research, as indicated in Table 5.2.

Table 5.2: Philosophical assumptions of the positivist, interpretive and design research perspectives (based on Vaishnavi & Kuechler, 2009).

Basic Belief	Positivist	Interpretive	Design
Ontology	A single reality, knowable, probabilistic	Multiple realities, socially constructed	Multiple, contextually-situated alternative world states, socio-technologically enabled
Epistemology	Objective, dispassionate, detached observer of truth	Subjective, values and knowledge emerge from researcher-participant interaction	Knowing through construction within a context, iterative circumscription reveals meaning
Axiology	Universal truth	Understanding: situated and description	Control: improvement and understanding
Methodology	Observation, quantitative, statistical	Participation, qualitative, hermeneutical, dialectical	Developmental, impacts by artefacts

In terms of ontology and epistemology, Vaishnavi and Kuechler state that in design research, such viewpoints shift as a study progresses though its various cycles. Hevner and Chatterjee (2010), however, criticise this view by pointing out that it is in fact the researcher's knowledge of the world that changes during a design research study and not assumptions about how the world is constructed. Axiologically, the design researcher values the control of the environment in search of understanding. It was mentioned in Section 5.4 that there is ongoing debate as to whether or not theory is an output of design research. However, Gregor and Hevner (2013) believe that even a partial or incomplete theory can be a valuable design research contribution if it provides a basis for further exploration.

Regarding the iterative and interventionist nature of design research, Vaisnavi and Kuechler (2009:20) conclude that the philosophical perspective of the design researcher is "very similar to the action research methodology of the interpretive paradigm". An interpretive-pragmatic view of DBR is supported by Barab and Squire (2004), Juuti and Lavonen (2006), Anderson and Shattuck (2012), and Hogue (2013). Barab and Squire (2004:1) suggest that DBR has a pragmatic philosophical underpinning, in which "the value of a theory lies in its ability to produce changes in the world".

5.5.3. Design-based research as research methodology

Design-based research is systematic and flexible, and aimed at improving educational practices (MacDonald, 2008). DBR methods attempt to bridge theory and practice in education as they uncover relationships between educational theory, designed artefacts, and practice (Design-based Research Collective, 2003). DBR is an approach that can improve educational research to yield discernable benefits and impact on practitioners, while also being more socially responsible (Reeves, Herrington & Oliver, 2005).

According to Barab and Squire (2004) and Amiel and Reeves (2008), DBR is fundamentally a problem-solving paradigm, also described as "complex interventionist research" (Bell, 2004:243). Furthermore, DBR can employ quantitative or qualitative research methods in a flexible way, as the researcher can adjust data collection methods in response to emerging research questions and to address research goals (Collins, Joseph & Bielaczyc, 2004; Hoadley, 2004). "Design-based research integrates the development of solutions to practical problems in learning environments with the identification of reusable design principles" (Herrington, McKenney, Reeves & Oliver, 2007:4089).

The design-based researcher aims at making both practical and scientific contributions by designing and creating effective interventions or authentic artefacts to solve complex, substantial problems in the real world where direct application of theory is not sufficient to solve those particular problems (De Villiers, 2012). According to Hay, Kim and Roy (2005:34), DBR holds much potential for "the development of emergent technology for learning where new technology affordances are explored and developed in a principled fashion". Therefore, DBR is an important approach for understanding why, how and when educational innovations work in practice.

The DBR process differs from experimental research in a number of ways:

- DBR involves real-world situations that contain limitations, complexities, and dynamics, while laboratory experiments are conducted in the laboratory in a controlled environment.
- In DBR, researchers try to characterise a complex situation through iterative and flexible revisions of the research design. In contrast, laboratory experiments are usually focused on a single dependent variable. The design-based researcher frequently follows new revelations where they lead, modifying the intervention as the research progresses (Hoadley, 2004).
- DBR is conducted in the real-world context and involves social interactions. Researchers conducting laboratory experiments attempt to isolate participants to prevent them from interacting with the outside world.
- DBR investigates educational problems by developing designs in practice, whereas most experimental research studies test hypotheses.

- In a DBR study the participants' expertise impacts on decisions in the different phases of the research process, while in experimental research, the researchers are the decision makers throughout the entire research process. (Collins *et al.*, 2004)

Not only can DBR be distinguished from experimental research, as described in the previous paragraph, but it can also be differentiated from pure design. According to Edelson (2002), DBR can be distinguished from design in that DBR is theory-based, has clear research goals and produces empirical results. Iterative formative evaluation can be utilised to identify gaps between the current design and the ideal design goals, leading to researchers tweaking the design to meet the research goals.

These descriptions of DBR indicate that DBR is particularly suitable for this research study on improving training using virtual reality technology, as the study addresses a real-world complex problem which leads to dual outcomes in the forms of authentic artefacts and a theoretical evaluation framework. The next section outlines the characteristics of DBR, followed by a summary in Table 5.3 of the features relevant to this study.

5.5.4. Characteristics of design-based research

This section discusses distinguishing features of DBR. Various authors elaborate on the processes and details of DBR. The concepts are presented serially according to topics in the literature, followed by a comprehensive tabulated synthesis.

Dual goals

The Design-based Research Collective (2003) describes DBR as having the dual goals of designing learning environments as well as developing theories. Wang and Hannafin (2005:8) state that DBR "refines both theory and practice". The research should lead to sharable theories that help communicate relevant implications to practitioners and other designers. According to Sandoval (2004), the instructional intervention is the focus of the research. Such interventions can also be conceptual artefacts designed to explore and illustrate research issues related to solving practical problems (Marchand & Walker, 2009).

The value of DBR should be measured by its ability to improve educational practice (Design-based Research Collective, 2006). DBR investigates how learning and performance are supported through convergence of theory and innovative learning or training environments. Major purposes of DBR are:

- the development of design principles through systematic inquiry into the processes of teaching and learning,
- the development of methods, technologies and innovative tools, by which these design principles are put in practice (Reeves, West & Orrill, 2006), and
- the construction of theoretical frameworks that inform future designs (Bowler & Large, 2008).

As stated, DBR can produce both theoretical and practical interventions as its outcomes. Edelson (2002) proposed three types of theoretical outcomes:

- Domain Theories that describe learning situations involving students, teachers, learning environments and their interactions.
- Design Frameworks that provide a set of context-specific design guidelines.
- Design Methodologies that serve as guidelines on how to implement designs.

Complex problems in authentic settings

Dede (2005) claims that many general research studies have very limited practical application, as they focus on documenting statistically significant outcomes for trivial problems with low effect sizes. In contrast, DBR deals with important issues, sizeable effects and statistically significant results. According to Shavelson, Phillips, Towne and Feuer (2003), DBR in educational settings traces the evolution of learning in complex environments and produces effective instructional tools that address challenges in practice. Joseph (2004) states that DBR interventions are used to provide insight into learning in real-world contexts.

Wang and Hannafin (2005) argue that DBR is grounded in theory and the real-world context. Theory is both the foundation and the outcome. In addition, DBR is conducted by researchers in collaboration with industry practitioners in complex, authentic contexts.

DBR research should account for how designs function in authentic workplace settings. Methods are followed to document and connect processes of enactment to outcomes of interest (DBRC, 2003). Reeves (2006) outlines three cornerstone principles of DBR:

- DBR involves collaboration between researchers and practitioners on complex problems in authentic settings,

- Solutions to complex problems are proposed by integrating technological advances with known and hypothetical design principles, and
- Rigorous research is required to evaluate and refine designs and new design principles.

Interventionist and integrative

According to the Design-based Research Special Interest Group (DBRSIG, 2006), DBR is interventionist as it serves as a change agent, so that a purposeful change can be made in a functioning educational environment. Van den Akker, Gravemeijer, McKenney and Nieveen (2006) identify five characteristics of DBR. They suggest that it is interventionist, iterative, process-oriented, utility-oriented, and theory-oriented.

Wang and Hannafin (2005) suggest that, depending on the requirements of the research, both qualitative and quantitative research methods can be used to build up evidence to support the theoretical principles underlying a specific intervention. They also argue that DBR results are connected both to the design process through which results are generated and to the specific context and environment where the research is conducted. After results are obtained, researchers should provide guidance on how the findings can be applied and adapted to other contexts.

Collaborative and participatory

DBR is participative in that it involves engagement by practitioners and researchers in long-term collaborations (Reeves, Herrington & Oliver, 2004). Through such collaboration an iterative process of analysis, design, development and implementation is followed, resulting in contextually-sensitive design principles and theories (Wang & Hannafin, 2005). The Design-based Research Special Interest Group supports this notion by stating that researchers should collaborate with practitioners (DBRSIG, 2006).

Barab and Squire (2004:3) also point out that "participants are not 'subjects' assigned to treatments but instead are treated as co-participants in both the design and even the analysis".

Pragmatic yet theoretical

DBR is pragmatic as it not only solves current real-world problems by designing and enacting interventions, but also extends theories and refines design principles to eventually lead to substantial change in educational practice (Barab, Sadler, Heiselt, Hickey & Zuiker, 2007; Wang & Hannafin, 2005). The Design-based Research Special Interest Group advocate that theory should be used to design the intervention (DBRSIG, 2006).

Bowler and Large (2008) add further characteristics of DBR, some of which confirm the above:

- DBR is multi-purpose in that it serves theory, design and practice.
- Secondary research questions can evolve or emerge during the research process, that is, the research design and process are not rigid.
- The application of DBR creates emerging theory.
- The result of DBR is a working artefact.

Iterative and flexible

Due to the iterative design process in DBR, researchers can continuously refine design interventions to make them more applicable to practice. Amiel and Reeves (2008:35) indicate that "the development of design principles will undergo a series of testing and refinement cycles".

DBR requires interactive collaboration between researchers and practitioners as theories and interventions are continuously developed and refined through an iterative design process (Wang & Hannafin, 2005). This recursive process flows from analysis to design to evaluation and redesign, allowing more flexibility than traditional experimental approaches (Bannan-Ritland, 2003).

Table 5.3 summarises DBR features, along with references to the literature arranged chronologically, and a short description of each feature. The features have been extracted from the literature review in this section, and have been synthesised and classified by the present researcher into a consolidated presentation. Most of the authors referenced are cited in this chapter, but certain others are incorporated as confirmatory sources. This table is revisited in Chapter Ten, where it is indicated how each feature is applied in this study.

Table 5.3: Summary of DBR features synthesised by the researcher.

Characteristic	Elaboration	References
Appropriate for complex environments	DBR deals with important issues, sizeable effects and significant results in complex environments.	Shavelson *et al.*(2003), Bell (2004), Dede (2005), Wang and Hannafin (2005), Plomp (2007), Anderson and Shattuck (2012).
Problem-solving paradigm	DBR is fundamentally a problem-solving paradigm. It explores research issues related to solving real-world practical problems. It is solution-oriented and the solutions that are developed must be relevant to authentic organisational issues.	Hevner *et al.* (2004), Amiel and Reeves (2008), Marchand and Walker (2009), De Villiers (2012), Gregor and Hevner (2013).
Grounded in theory	Theory is the foundation and an outcome. Theory is used to design the initial intervention. Theoretical frameworks are built that inform future designs. This leads to theory building as the artefact exposes relationships between its elements.	Rossi and Sein (2003), Wang and Hannafin (2005), DBRC (2006), DBRSIG (2006), Van den Akker (2006), Barab *et al*. (2007), Herrington et al. (2007), Bowler and Large (2008), Gregor and Hevner (2013).
Collaborative and participative	Practitioners and researchers are engaged in long-term collaborations. The expertise of practitioners and researchers impacts on decisions in the different phases of the research process.	Collins *et al.* (2004), Reeves *et al.* (2004), DBRSIG (2006), Reeves (2006), Anderson and Shattuck (2012).

Flexible and adaptable	Theories and interventions are continuously developed and refined; the researcher frequently follows new revelations where they lead, modifying the intervention as the research progresses.	Hoadley (2004), Wang and Hannafin (2005), Bowler and Large (2008), MacDonald (2008).
Context-sensitive	Designs function in authentic settings. Results are connected to both the design process and the context where the research is conducted.	Zaritsky *et al.* (2003), Barab and Squire (2004), Wang and Hannafin (2005), Reeves (2006), Plomp (2007), Amiel and Reeves (2008), De Villiers and Harpur (2013).
Integrative	The relationships between theory, designed artefact, and practice are uncovered. There is a convergence of theory and innovative learning environments.	DBRC (2003), Amiel and Reeves (2008), Bowler and Large (2008), De Villiers and Harpur (2013).
Innovative	DBR addresses complex unsolved problems in unique or innovative ways. The development of solutions is informed by existing design principles and technological innovations. Methods, technologies and innovative tools are developed whereby new design principles can be put in practice.	Hevner *et al.* (2004), De Villiers (2005b), Wang and Hannafin (2005), Reeves *et al.* (2006), Amiel and Reeves (2008).
Iterative	A systematic and iterative process of analysis, design, development and implementation is followed by researchers to continuously refine design interventions to make them more applicable to practice.	Bannan-Ritland (2003), Zaritsky *et al.* (2003), Hoadley (2004), Wang and Hannafin (2005), Van den Akker *et al.* (2006), McKenney and Reeves (2012), De Villiers and Harpur (2013).

Dual outcomes	The outcomes of DBR are: (i) a practical contribution in the form of an innovative product or intervention; (ii) a set of design principles or guidelines, with the objective of a theoretical contribution.	Barab and Squire (2004), Wang and Hannafin (2005), Amiel and Reeves (2008), De Villiers (2012), Teräs and Herrington (2014).
Pragmatic, yet theoretical	DBR is aimed at addressing actual challenges in organisations as it extends theories and refines design principles to lead eventually to substantial change in practice. The theory should be transferable to other contexts.	Hay *et al.* (2004), De Villiers (2005b), Wang and Hannafin (2005), Reeves *et al.* (2006), Barab *et al*. (2007), Amiel and Reeves (2008), MacDonald (2008).
Artefacts	Artefacts that are authentic, tangible products are produced as purposeful practical interventions in a functional environment.	Bell (2004), De Villiers (2005b), Van den Akker *et al.* (2006), Herrington et al. (2007), Bowler and Large (2008), Marchand and Walker (2009), Gregor and Hevner (2013).
Evaluation	DBR involves rigorous and reflective inquiry to test and refine artefacts. Design theories and design principles are evaluated and refined. The quality, utility and efficacy of the designed artefacts are also evaluated.	Hevner *et al.* (2004), Reeves (2006), Amiel and Reeves (2008), Gregor and Hevner (2013).
Mixed-methods	DBR employs quantitative or qualitative research methods, or both, depending on the nature of the particular research being undertaken.	Collins *et al.* (2004), Hoadley (2004), Wang and Hannafin (2005), Venable (2010).

5.5.5. Principles for application of design-based research

DBR is a purposeful and systemic methodology that requires rigorous, disciplined and iterative inquiry in order to generate credible, practical and contextual design theories. Wang and Hannafin (2005) propose nine principles for applying DBR in technology-enhanced learning environments.

1. Support design with research: Identify resources from literature from the outset.
2. Set practical goals: Setting goals helps to enhance rigour and enforce discipline.
3. Conduct research in workplace settings: Analysis of literature and the real-world setting is required for innovative design in practice.
4. Collaborate with participants: Researchers need to become familiar with the people, resources and constraints of the research environment.
5. Combine research methods: Qualitative and quantitative techniques are often both employed in DBR.
6. Analyse data continuously: In order to improve the design and to address theory generation, data should be analysed immediately and also retrospectively.
7. Refine designs continuously: Intermediate design goals can be reached by refining designs iteratively. Such refinements can also include new innovations, and can be due to external or unanticipated influences.
8. Document design principles: Context-sensitive design principles should be specified to inform future practice.
9. Validate the generalisability of the design: The effectiveness of a DBR design is measured both from a perspective of how successfully local needs are addressed as well as the applicability of the design principles to other workplace settings.

5.6. Research design of this study: a design-based research model

This section presents the research processes used in this study and explains how a customised process model was constructed by referring to preceding research models of the design and development research genre. The content of Section 5.6 and some of the figures in the section, were included in a conference paper by Van Wyk and De Villiers (2014), presented by the researcher at the South African Institute of Computer Scientists and Information Technologists' annual conference in 2014 and published in the conference proceedings. The paper is attached as Appendix D-3. The work in the paper was done specifically for the purpose of this doctoral study. It culminates in a DBR process model defined by the researcher as a foundation for this research, which is presented in Figure 5.4.

Wang and Hannafin (2005) compared certain DBR variants and differing terminology in the literature, some of which have been mentioned in this chapter. These include:

- design experiments (Brown, 1992; Collins, 1992),
- developmental research (Richey & Nelson, 1996),
- design research (Cobb, 2001; Collins *et al.*, 2004; Edelson, 2002; Reeves, 2006),
- formative research (Reigeluth & Frick, 1999), and
- development research (Reeves, 2000; Richey, Klein & Nelson, 2003; Van den Akker, 1999).

They conclude that even though each variant has a slightly different focus, the underlying goals and approaches are similar.

With this in mind, and since few process models for DBR exist in the literature, the present researcher proposes a model for DBR which is primarily based on its immediate predecessor, development research. The proposed model is also influenced by the processes of design science research as described by Vaishnavi and Kuechler (2009), and the DBR cycle as explained by Amiel and Reeves (2008).

Before presenting the proposed model in Subsection 5.6.5, these three predecessors are addressed in Subsections 5.6.1, 5.6.2 and 5.6.3 respectively, after which a consolidated view of these predecessors is presented in Subsection 5.6.4.

5.6.1. Development research

One of the variants from which DBR evolved is the development research (DR) methodology (Van Den Akker, 1999; Reeves, 2000). A development research approach was used when complex learning content, created to function in complex real-world contexts, required research designs that assessed the development process as well as the outcome of the intervention. It was performed in order to optimise and gain a sound basis for development activities, primarily in the context of educational technology applications. DR is a problem-oriented and interdisciplinary research methodology that is characterised by the development of prototypical products and the generation of methodological directions for design and evaluation of such products (Van den Akker & Plomp, 1993).

The process in DR is based on the dynamic and flexible ADDIE design model that originated in instructional technology. The five phases — Analysis, Design, Development, Implementation, and Evaluation — represent a dynamic, flexible guideline for building effective training and performance support tools (Molenda, 2003). Within DR, the ADDIE process is followed in an iterative manner, where the results of each formative evaluation are used for analysis and further development in the next cycle (De Villiers, 2012).

DR has a dual focus. It is applied for the development of practical and innovative ways of solving real problems, and it also proposes general design principles to inform future decisions. DR searches for "new and innovative solutions, while also seeking findings that are transferable, practical and socially responsible" (De Villiers, 2005b). According to Plomp (2000), DR is aimed at:

- reducing uncertainty of design decisions,
- generating concrete recommendations for quality improvement,
- testing general design principles, and
- stimulating professional development.

The DR model used by Plomp (2002) and Van den Akker (1999) refers to two types of outcomes of an intervention: *immediate outcomes* relate to results obtained by using an intervention or product within the cyclic process, and *distant outcomes* emerge in the form of generalisable principles as results of the process. These two outcomes represent the dual focus.

5.6.2. The design science research methodology

Another variant is the general methodology of design science research in the information systems discipline, which was introduced in Section 5.4. DSR involves "analysis of the use and performance of designed artefacts to understand, explain and very frequently to improve on the behavior of aspects of information systems" (Vaishnavi & Kuechler, 2009:2).

According to Vaishnavi and Kuechler (2009), the following are the process steps of a typical DSR project.

1. Problem Awareness: awareness can emerge from literature or new developments in an industry or discipline, and leads to a proposal.
2. Suggestion: new functionality is envisioned and a tentative design is generated.
3. Development: the tentative design is implemented, producing an artefact.
4. Evaluation: the artefact is evaluated according to the criteria specified in the proposal.

5. Conclusion: the project is considered complete when results are adjudged as satisfactory.

As described in Section 5.4, the outcomes of DSR may be one or more of the following artefacts: construct, model, method or instantiation (Gregor, 2002; March & Smith, 1995; March & Storey, 2008).

5.6.3. The design-based research cycle

The DBR cycle is related to the preceding models and, in particular, has evolved from development research. As explained in Amiel and Reeves (2008), it comprises the following steps:

1. Analysis of practical problems by researchers in collaboration with practitioners.
2. Development of solutions informed by existing design principles and technological innovations.
3. Iterative cycles of testing and refinement of solutions in practice.
4. Reflection to produce design principles and to enhance the implementation of the solution.

Each step can lead to refinement of previous steps and processes. Amiel and Reeves state that the outcomes of DBR are a set of design principles or guidelines that can be implemented by other researchers working in similar contexts, with the ultimate objective being the development of theory, as well as a practical contribution in the form of an innovative product or intervention. This may only be achieved after long-term engagement and several cycles of specific design interventions. Ma and Harmon (2009) expanded on the work of Amiel and Reeves by elucidating the process for one iteration of the DBR cycle to indicate, in detail, the steps involved. These steps are shown in Figure 5.2.

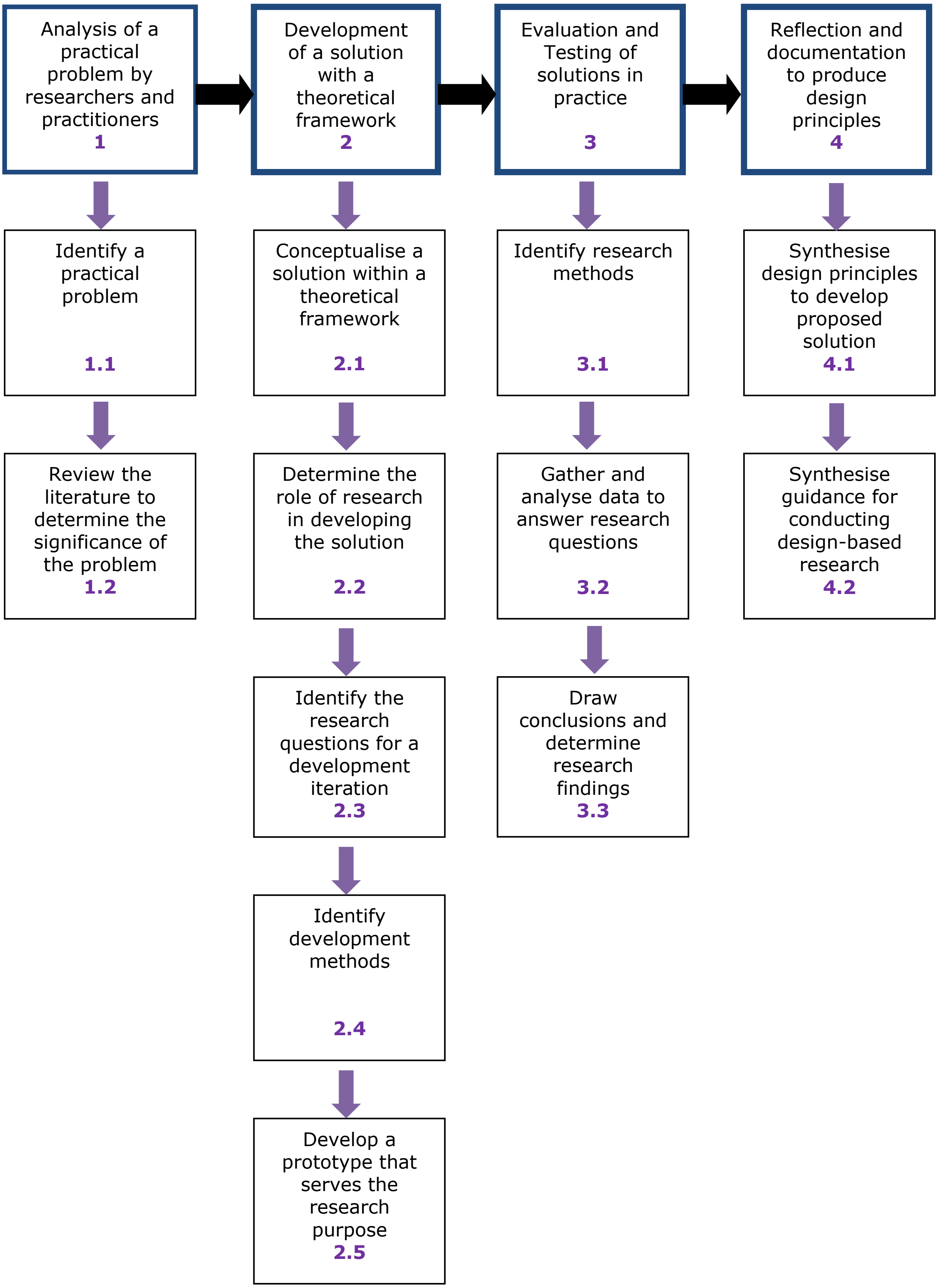


Figure 5.2: Design-based research: the process for one iteration (Ma & Harmon, 2009:77).

5.6.4. Consolidation of the foregoing models

Figure 5.3 is a composite diagram depicting:

- the synthesised development research model (De Villiers, 2005b; Reeves, 2000; Van Den Akker, 1999);
- the general methodology of design science research (Vaishnavi & Kuechler, 2009); and
- the design-based research cycle (Amiel & Reeves, 2008).

It is evident from the diagram that there are major similarities in the phases of the three presented procedures. Each procedure starts with a problem analysis phase. In the case of DR and DBR there are searches for innovative ways of solving real problems, while a DSR project also analyses problems related to new developments in a discipline and includes a formal proposal. The next phase in DR and DSR is to focus on producing a tentative design of a potential solution that is developed in the subsequent phase. In the DBR cycle, the development phase also includes the design of a new, authentic innovation. The developed solution is then implemented and evaluated, which in all three cases is done in real-world practice.

All three procedures in Figure 5.3 are iterative, in that feedback from a certain phase can lead to improved input into previous phases, resulting in redesign, refinement and improvement. A DSR project concludes when the evaluated results are acceptable when measured against criteria specified in the project proposal. The iterations of the DR process produce an implemented solution in a real context, for example, in industry, as well as generic design principles, an output similar to that of the reflection phase of DBR. DBR focuses on producing solutions that can also be adopted elsewhere, and on the development of new theory to guide similar research and development.

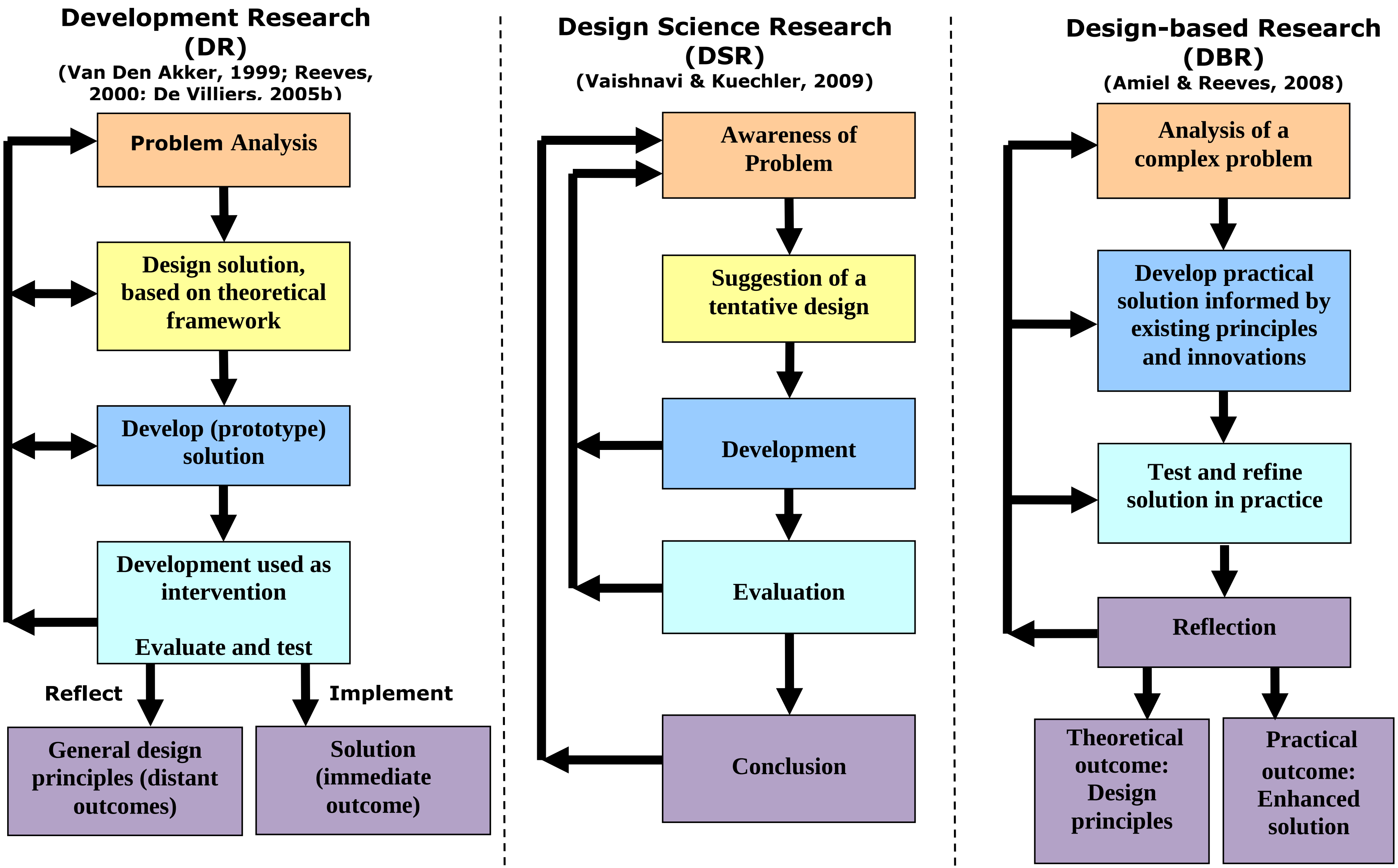


Figure 5.3: Composite diagram depicting the Development Research Model, the General Methodology of Design Science Research, and the Design-based Research Cycle.

5.6.5. A new DBR model

Figure 5.4 depicts a new generic model of DBR synthesised by the present researcher, and influenced by the procedures in the three subdiagrams of Figure 5.3, namely the DR model, the general DSR methodology used in the information systems domain, and the DBR cycle. It also applies the process for one DBR iteration as depicted by Ma and Harmon (2009) and shown in Figure 5.2. The terms in red in Figure 5.2 emphasise key characteristics of DBR, as discussed in Section 5.5.4.

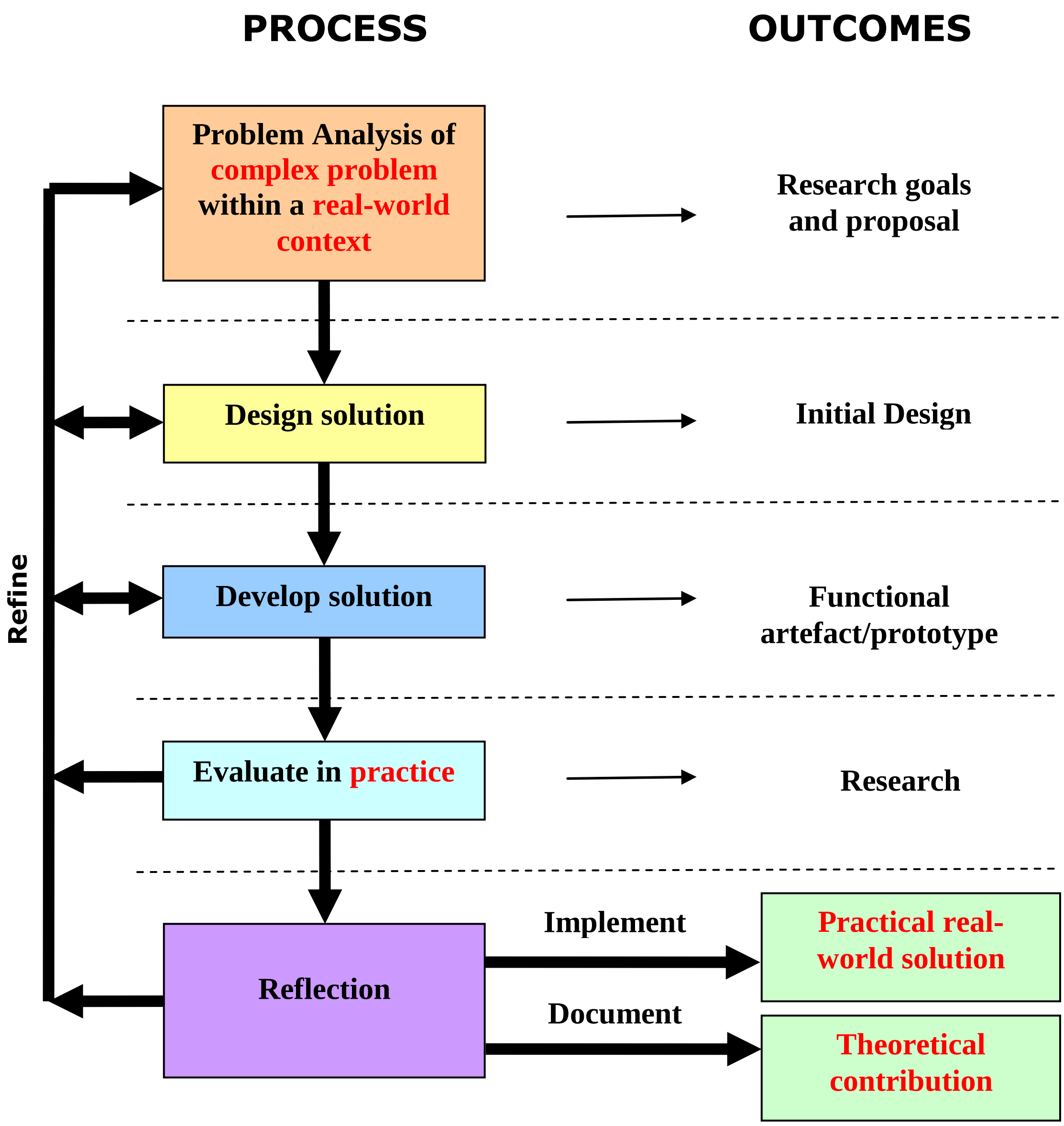


Figure 5.4: A synthesised generic model for design-based research.

5.6.5.1. Characteristics of the new DBR model

Figure 5.4, which presents the new DBR model synthesised by the researcher, integrates the phases and processes of the precedents, as well as explicitly indicates the outcomes associated with each phase on the right-hand side.

1. Problem analysis within context: A practical, authentic problem is identified in a complex environment and literature is reviewed to determine the significance of the problem and to identify current theory on the topic. Researchers and practitioners collaborate in analysing the problem and establishing research goals. The outcome of this step is an explicit research proposal containing research goals.
2. Design solution: An initial design is proposed to address the problem. The appropriateness of the design is influenced by contextual limitations and the complexity of interactions that occur in real-world settings.
3. Develop solution: A prototype is developed that serves the research purpose. Development is informed by existing design principles and technological innovations. The outcome is an innovative, functional artefact (construct, model, method or instantiation), aspects imported from the rigorous realm of design science research.
4. Evaluate in practice: The artefact is tested in a real-world setting. Data is collected and analysed to answer the research questions and to construct principles or theory.
5. Reflection, leading to dual outcomes:
   - Practical real-world contribution – Reflection enhances the implementation of the solution. As reflection occurs upon the data, new designs can be developed and implemented, leading to an ongoing subcycle of design-reflection-design.
   - Theoretical contribution – Design principles should be continuously and cumulatively documented in order to be transferable and utilised by others in similar settings. New theory may be developed, but may require multiple DBR cycles over a long term to develop sufficiently.

This new model has certain distinct attributes as well as the standard features of DBR. The outcomes of each step are specified, notably an explicit research proposal as an outcome of problem analysis. The arrows in the diagram, some uni-directional and others bi-directional, emphasise the iterative nature of the process. Feedback to previous steps is labelled *Refine* since the focus is not only on evaluating the artefact, but also on systematically refining the innovation. This view of DBR includes evolution of the innovation or product that is the designed solution, but also refinement of the problem, the methods and frameworks, the tools used in design and evaluation, and the design principles. The process culminates in dual outputs: (i) an implemented solution that addresses the original problem in its real-world setting and (ii) documented design principles and/or other theory that can guide similar research and development efforts.

The colours indicating each process in Figure 5.4 are re-used as consistent colour coding in subsequent diagrams.

5.6.5.2. Comparison of new DBR model with Amiel and Reeves' DBR model

The new DBR model, shown in Figure 5.4, differs from the DBR model of Amiel and Reeves as follows:

- It extends the classic DBR model by including the design of solutions which are not necessarily based on "existing design principles and technological innovations" (Amiel & Reeves, 2008:34) nor "within a theoretical framework" (Ma & Harmon, 2009:78) nor "drawn from the existing knowledge/theory base for the problem area" (Vaishnavi & Kuechler, 2009:10). In the domain of virtual reality systems, where this work resides, design theory is relatively new and established design principles are not available. The new model allows for conceptualisation of solutions beyond existing mature theoretical frameworks, due to the innovative nature of the technology being applied.
- The new model applies a feature of DSR by advocating a proposal as an output of the first phase. Amiel and Reeves (2008:5) describe the first phase of DBR (problem analysis) as the "negotiation of research goals between practitioners and researchers", but Vaishnavi and Kuechler (2009:7) describe the output of the first phase of DSR (problem awareness) as a "proposal, formal or informal, for a new research project". Their formal proposal includes a tentative design and performance criteria to evaluate the prototype. Similarly, this new DBR model requires a proposal that includes research goals.
- The new model adapted the DBR model to include a theoretical outcome that is not merely a set of design principles. Amiel and Reeves (2008:35) describe the outcomes of DBR as "a set of design principles or guidelines derived empirically and richly described, which can be implemented by others interested in studying similar settings". Being a design-based research methodology, the importance of design principles as an output is indeed acknowledged, but provision is also made for new theoretical contributions that extend even further. Such contributions, importantly, should inform future design and evaluation in similar environments in practice.

5.6.6. Research process flow

The research design followed in this study is based on the synthesised DBR model depicted in Figure 5.4. To explain how the model was applied in the context of this study, Figure 5.5 shows the process flow from problem to solution, where each DBR cycle, indicated as a blue circle in the diagram, is an instance of the full DBR model shown in Figure 5.4. In Figure 5.5 the red blocks indicate actions, the blue ovals represent instances of the DBR model and the green blocks indicate the specific artefacts or theory deliverables that are the outcomes of the process in this research.

Cycle 1

The research process commenced with a definition of the real-world practical problem. Every year more than a hundred workers die in the South African mining industry and thousands are injured (Webber-Youngman & Van Wyk, 2013). As stated in Chapter One, the aim of this study is to propose, model, prototype and evaluate two novel electronic training interventions to improve the safety of mine workers. These e-training systems, implemented by VR technology to simulate underground conditions and potential hazards, will supplement conventional classroom training.

During DBR Cycle 1, a desktop VR training prototype, *Look, Stop and Fix* (*LSF*), was designed and developed based on the problem analysis and preceding literature reviews. This prototype simulates the underground working areas, incorporating potential hazards. Mine workers need to spot the hazards, identify them correctly, and indicate appropriate actions to be followed in response to each hazard. Failure to correctly identify a hazard or to specify the correct action to deal with such a hazard, causes an animation to play out, displaying the possible disastrous consequences of ignoring or incorrectly responding to such a hazard. The *LSF* prototype was evaluated in practice by user surveys at the Mine. Details of the evaluation are given later in this section. After the evaluation, the reflection step identified several problems that should be addressed.

Evaluation framework and Cycle 2

Following DBR Cycle 1, a formal evaluation framework and a set of criteria were developed for evaluating desktop VR training applications in the next round of evaluation, which involved heuristic evaluation by experts. During DBR Cycle 2, the prototype design was refined and then evaluated using this evaluation framework. This cycle also included several internal subcycles of the DBR steps, resulting in an improved version of *LSF.*

Case study and Cycle 3

Upon reflection, the Mine indicated that the prototype, which focuses on generic hazards, should be expanded to focus on the major causes of incidents at the Mine. This led to a case study which investigated and identified the causes of incidents, culminating in the design and development of a new geological conditions prototype in DBR Cycle 3. This prototype was called *Interactive Simulated Geological Conditions* (*ISGC*). In the *ISGC* prototype, trainees have to identify 21 different geological conditions occurring in that particular mine and specify the associated risks and controls for each condition. Animations are shown of the possible results of ignoring or not correctly addressing the geological conditions.

The evaluation framework was again used to evaluate this new follow-up prototype and problematic issues in *ISGC* were identified. Details of all the evaluations are given later in this section.

Improvement of prototypes and Cycle 4

In the process of evaluating both the prototypes, namely *LSF* and *ISGC,* inadequacies also emerged in the evaluation framework. After implementing improvements to both prototypes, a further DBR cycle (Cycle 4) was applied to improve the evaluation framework itself, so as to strengthen future evaluations. A meta-evaluation questionnaire was designed and a meta-evaluation of the evaluation framework identified possible improvements.

Outcomes of DBR process

Finally, Figure 5.5 also indicates the outcomes resulting from the processes as shown. The solution consisted of dual outcomes in line with design-based research, involving a practical real-world contribution in the form of novel desktop VR training systems and a theoretical contribution which is an evaluation framework for evaluating desktop VR training systems.

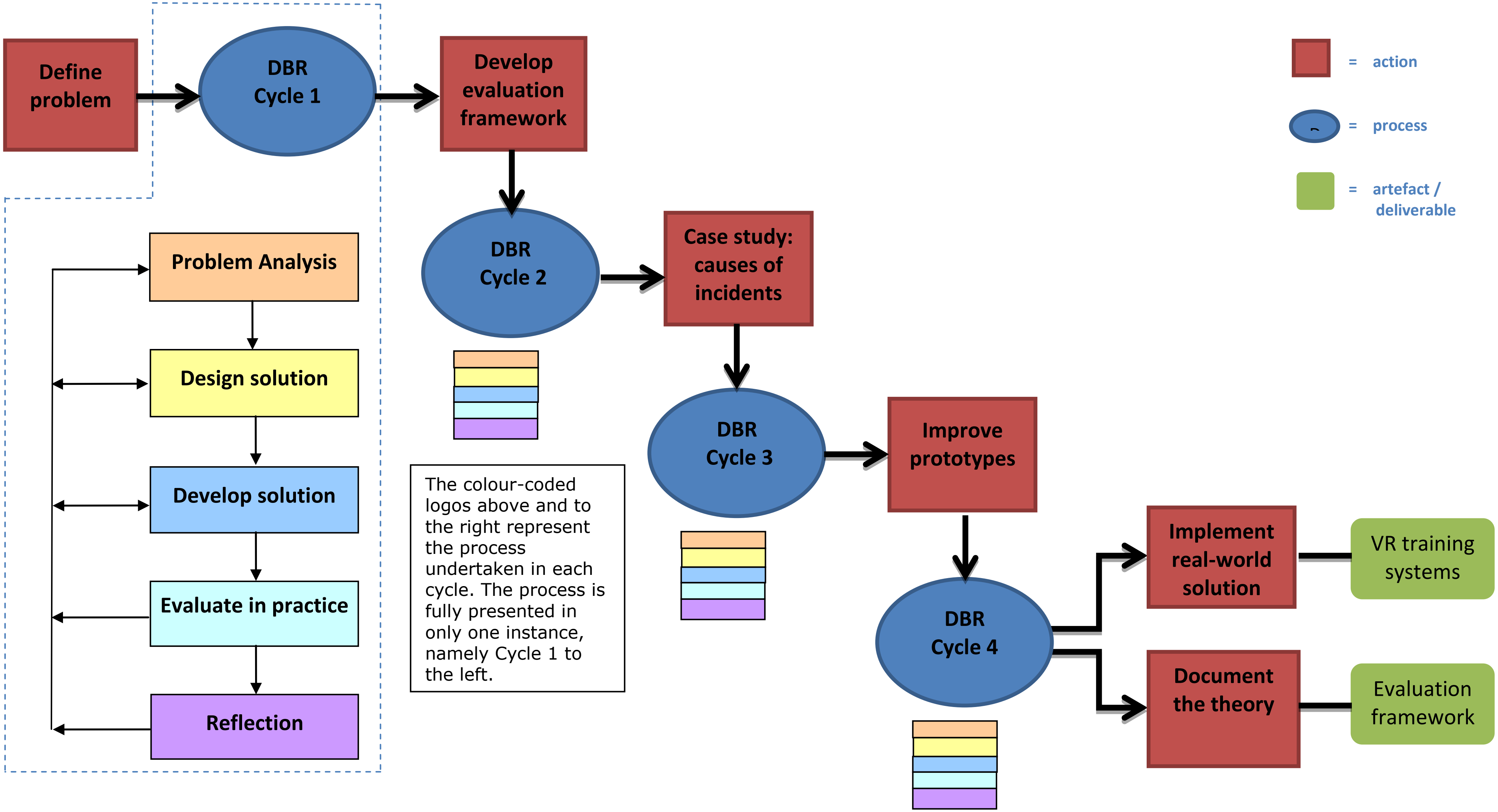


Figure 5.5: Process flow diagram of this research study.

5.6.7. Detailed process flow

Although four cycles of the DBR process were executed, each cycle is not presented as a separate study in this thesis. The chapters following this research methodology chapter thus do not correspond one-on-one to associated cycles of the DBR research process flow, but rather focus on important aspects of the process as a whole, which led to the dual outcomes of this research. These aspects are respectively: a usability context analysis (Chapter Six); the design of the prototypes (Chapter Seven); the evaluation of the prototypes (Chapter Eight); and the meta-evaluation of the evaluation framework (Chapter Nine). For this reason, elaborated descriptions of each step in the four cycles are included in this chapter, in order to overview the processes followed and the steps taken.

Figure 5.6, an elaboration of Figure 5.5, provides more detailed information on the four DBR cycles in Figure 5.5. It graphically represents the discussion above by incorporating detail to explain how the steps of each DBR cycle in Figure 5.5 were executed in line with the generic DBR model in Figure 5.4. Each cycle in Figure 5.6 is situated within a frame. The colour coding used in Figure 5.6 to indicate the steps in each cycle is the same as the colours in Figures 5.4 and 5.5.

The blue arrows in Figure 5.6 indicate the links between the cycles.

- The results of the reflection step in Cycle 1 form the input to the problem analysis step of Cycle 2,
- The results of the reflection step in Cycle 2 form the input to the problem analysis step of Cycle 3, and
- Due to the fact that the same evaluation framework was used to evaluate both the *LSF* and *ISGC* prototypes in Cycles 2 and 3 respectively, the results of the reflection steps in Cycle 2 and Cycle 3 both inform the problem analysis step for Cycle 4, where a meta-evaluation is performed, leading to an improved evaluation framework.

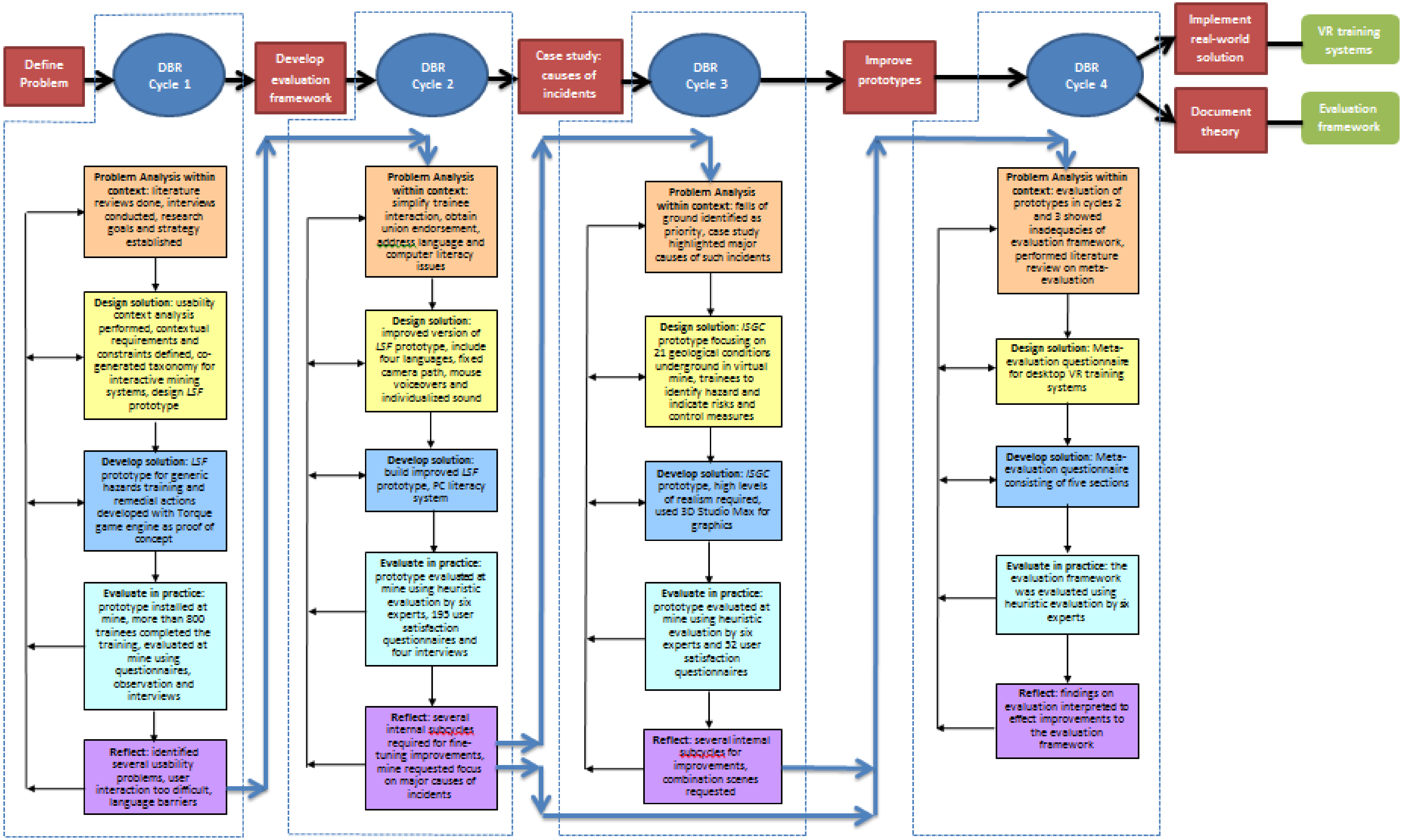


Figure 5.6: Research process flow with summary details of each DBR cycle.

The next subsections detail the steps in the four cycles of the DBR model in Figure 5.6. To assist recall, the definition of each step in the model, as given in Section 5.6.5, is repeated below in the description of the first cycle (in blue font), followed by the associated application in this study. For the other three cycles the definitions are not given again, but the explanations are provided. After the explanations in each cycle, a table summarises the research methods used in each cycle. Four cycles of the model were executed.

DBR Cycle 1

1. **Problem analysis within context**:

- A practical, authentic problem is identified and literature is reviewed to determine the significance of the problem and current theory on the topic.

  The authentic problem of safety training in the mining industry was pinpointed, as explained in Chapter One, and relevant literature and documents were reviewed to determine the extent, impact and nature of the problem and to identify current theory on the topic. These literature reviews were discussed in Chapters Two, Three and Four.
- Researchers and practitioners collaborate in analysing the problem and establishing research goals.

  Interviews with three mine managers, two safety, health and environment officers and three mine training managers provided more information and confirmed the need for improved safety training.
- The outcome of this step is a proposal containing research goals.

  This led to the establishment of research goals and a research strategy for the study. An extensive research proposal was written and refined into Chapter One.

2. **Design solution**: An initial design is proposed to address the problem. The appropriateness of a design is influenced by the contextual limitations and the complexity of interactions that occur in real-world settings.

   Information obtained from the literature studies, as well as a usability context analysis (UCA), informed the initial design of a prototype to address the problem. The UCA (discussed in Chapter Six) indicated the contextual requirements and constraints for VR training systems for the mining industry (Van Wyk & De Villiers, 2008).

3. **Develop solution**:

- Develop a prototype that serves the research purpose. Development is informed by existing design principles and technological innovations.

  As a member of the International Group for Virtual Reality in Mining, the present researcher was involved in the development of a taxonomy of interactive computer-based visualisation systems and content for the mining industry (Stothard, Squelch, Van Wyk, Schofield, Fowle, Caris, Kizil & Schmid, 2008). A prototype, called *Look, Stop and Fix (LSF)*, was developed based on the information derived in Step 2 above, and on the technological options available for applying VR within the defined context, as indicated in the taxonomy. *LSF* simulates the underground working areas and focuses on general hazard recognition and remedial actions. The environment presents potential hazards; mine workers must identify them and indicate responsive actions. The researcher was assisted in the development and programming of *LSF* by students from the Centre for Creative Technologies at the Tshwane University of Technology.

- The outcome is an innovative, viable artefact.

  The prototype was a functional computer system presenting a 3D virtual underground environment, where the haulage and stope areas were realistically simulated. A total of 27 generic hazards were present in this environment. Trainees interact with the system to spot these potentially hazardous conditions, to identify the hazards correctly, and to indicate which action/s should be taken to address the situation.

4. **Evaluate in practice**:

- The artefact is tested in the real-world setting.

  The *LSF* prototype was installed at a training facility at No. 10 shaft at the Mine, where computers were acquired specifically for this purpose and housed in a venue in the administration building. All the workers at this shaft were trained using the prototype.

- Data is collected and analysed to answer the research questions and to construct design principles.

  Data collection was done through informal observation and interviews with trainees. More than 800 trainees completed the e-training. Questionnaires were completed by 221 trainees after completion of the prototype training and structured interviews were conducted with 23 randomly selected trainees.

5. **Reflection**: Reflection enhances solution implementation. As the researcher reflects on the data, new designs can be developed and implemented, which leads to a continuous

cycle of design-reflection-design. Design principles should be documented in order to be transferable and utilised by others in similar settings. New theory may be developed, but may require multiple DBR cycles over a long term.

Based on the data collected, various improvements were made to the prototype, which resulted in a new version of *LSF*. Following the general success of the training done at No. 10 shaft, as described in Cycle 1, Mine management decided that all employees who worked underground should be trained with the improved version of the *LSF* prototype. The mine trade union wanted to be involved in the decision and a group of four officials asked to view the prototype. After they had undergone the training, they endorsed the program strongly and approved its use at all shafts.

The aspects mentioned above, are summarised and consolidated in Table 5.4.

Table 5.4: Research Design: Cycle 1 research methods.

Cycle 1		**Outputs**
Participants	• 3 Mine Managers • 4 SHE Officers • 3 Mine Training Managers • >800 mine workers completed prototype training	• Taxonomy of interactive visualisation systems • Prototype: *LSF* (generic hazards identification and remedial actions), using the Torque game engine as the main development environment
Data Collection Methods	Literature studies, questionnaires and interviews	**Data**
Data Analysis	Quantitative – elementary statistical analysis Qualitative – noting patterns and themes, clustering, interpretive analysis	• Questionnaires completed by 221 trainees • Semi-structured interviews with 4 Union Officials • Transcripts of interviews with Mine Managers, SHE Officers and Mine Training Managers • Transcripts of interviews with 23 randomly-selected trainees

DBR Cycle 2

1. **Problem analysis within context**:

   Certain problems emerged from informal observation of trainees using the prototype at No. 10 shaft. The main issue that became evident, was the trainees' inexperience in using computers, which caused them to struggle in interacting with the system. A way had to be found to make user interaction simpler. A further problem was the language barrier, with many miners being unable to read English.

2. **Design solution**:

   An improved version of the *LSF* system was designed to simplify user interaction. Due to their lack of exposure to computers, many trainees struggled with the computer mouse. To provide pre-training intervention, a PC literacy course was designed, which focused mainly on mouse skills.

3. **Develop solution**:

   The improved version incorporated four languages (English, Tswana, Xhosa and Sepedi) with mouse voiceovers and individualised sound using headsets. Interaction was simplified by having a camera pan through the virtual environment on a fixed path, eliminating the requirement for trainees to manipulate themselves through the simulated underground environment.

4. **Evaluate in practice**:

   The Mine acquired more computers and the improved version of *LSF* was installed at the training centre. The system was evaluated using heuristic evaluation by experts, as well as by a user satisfaction survey entailing questionnaires completed by end-users. The heuristic evaluation was conducted by six expert evaluators. The user satisfaction questionnaire was completed by 195 trainees after completion of the *LSF* training. These participants were the trainees doing the *LSF* training on the dates that the researcher visited the mine to collect data. They comprised a typical sample. Results of the *LSF* evaluations are discussed in Chapter Eight, along with the evaluations of *ISGC* from Cycle 3.

5. **Reflection**:

   Due to the success of *LSF,* the Mine requested further extensions. A requirement was that the next system should focus not only on generic hazards, but specifically on the main causes of accidents.

Table 5.5 summarises the research methods applied in DBR Cycle 2.

Table 5.5: Research Design: Cycle 2 research methods.

Cycle 2		**Outputs**
Participants	• 4 Union Officials • 2 Usability Experts • 2 Mine Training Experts • 2 VR Development Experts • >16000 mine workers completed prototype training	• PC Literacy system • Prototype: Improved *LSF* (camera on fixed path, 4 languages, mouse voiceover), still using Torque game engine
Data Collection Methods	Interviews and Survey research Heuristic evaluation	**Data**
Data Analysis	Quantitative – statistical analysis Qualitative – tabulating data, organised by groups, calculating frequencies of occurrences and responses	• Heuristic evaluation by 6 experts • User satisfaction questionnaire completed by 195 end-users

DBR Cycle 3

1. **Problem analysis within context**:

   Due to falls of ground being the greatest contributor to mining injuries, a case study was done to analyse the circumstances relating to falls-of-ground incidents at the Mine. Details of the case study, which was an input to Cycle 3, are given in Chapter Seven.

2. **Design solution**:

   A prototype of a new system was designed with the particular aim of focusing on the major geological hazardous conditions contributing to falls of ground. This system was called *Interactive Simulated Geological Conditions (ISGC)* and 21 geological conditions were simulated in a virtual environment of the underground workplace. In *ISGC* the trainee is required to identify the condition correctly and specify the associated risks and control measures for each condition. Although *ISGC* was a new prototype, its design and development were strongly influenced by lessons learned from the design, evaluation and reflection of the *LSF* prototype in DBR Cycles 1 and 2. The initial target group for *ISGC* was 52 employees of the higher ranks of underground workers, that is, shift supervisors and mine overseers.

3. **Develop solution:**

   For accurate simulation of the geological conditions, graphics with a high level of realism were required. This could not be achieved with the Torque game engine environment used for the two versions of *LSF,* and the models and animation in *ISGC* were therefore developed in 3D Studio Max. Once again, the researcher was assisted in the development and programming by students from the Centre for Creative Technologies at the Tshwane University of Technology.

4. **Evaluate in practice**:

   The *ISGC* system was installed at the training centre and evaluated by heuristic evaluation and a user satisfaction questionnaire. The heuristic evaluation was done by the same six experts who evaluated the *LSF* prototype in Cycle 2. The user satisfaction questionnaire was completed by 52 trainees after completion of the *ISGC* training. These participants were the trainees doing the *ISGC* training on the dates that the researcher visited the mine to collect data. They comprised a typical sample. Results of the evaluation are discussed in Chapter Eight.

5. **Reflection**:

   Within Cycle 3, several internal design-reflection-design cycles led to many improvements to *ISGC*, including orientation labels, additional visual learning material, and scenes containing combinations of more than one geological hazard, as is practically experienced underground.

The research methods applied in DBR Cycle 3 are summarised in Table 5.6.

Subsequent refinements resulted in the *ISGC* system being installed as a formal training system at the Mine. All underground mine workers do this training on return from their annual leave. A trainee must obtain a pass score before being allowed to resume work underground. *ISGC* is, at the time of writing, used at the Mine as an implemented solution to improve safety training and is thus addressing the original problem in a real-world setting, which is in line with the spirit of DBR.

Table 5.6: Research Design: Cycle 3 research methods.

Cycle 3		**Outputs**
Participants	• 2 Usability Experts • 2 Mine Training Experts • 2 VR Development Experts • >200 mine workers completed prototype training	• Prototype: *ISGC* (geological hazards recognition and remedial actions), using 3D Studio Max • Improved levels of realism
Data Collection Methods	Case Study and Survey research Heuristic evaluation	**Data**
Data Analysis	Quantitative – statistical analysis Qualitative – tabulating data, comparing groups, calculating frequencies of responses, interpretive analysis, noting themes in open-ended responses	• Findings of case study to determine major causes of accidents • Heuristic evaluation by 6 experts • User satisfaction questionnaire completed by 52 end-users

DBR Cycle 4

1. **Problem analysis within context**:

   The evaluations of the *LSF* and *ISGC* prototypes during Cycles 2 and 3 not only provided valuable information regarding the prototypes, but also indicated that the *DEVREF* Evaluation Framework had inadequacies. It was the appropriate time to determine the effectiveness of the evaluation framework itself by performing a meta-evaluation on the instrument and criteria that had been used as the basis of the heuristic evaluations. The meta-evaluation and its findings are discussed in Chapter Nine.

Subsequent findings from the evaluations of the *LSF* and *ISGC* prototypes, plus problems and gaps pinpointed in the meta-evaluation, led to an improved version of the *DEVREF* Evaluation Framework, which is also presented in Chapter Nine.

2. **Design solution**:
   To design a meta-evaluation instrument, the researcher studied the literature on meta-evaluation of evaluation frameworks, revisited the literature used in generating the *DEVREF* heuristic evaluation framework, and consolidated the most important aspects into a set of evaluation statements. This is addressed in Section 9.2.

3. **Develop solution:**
   The evaluation statements were categorised in the form of a meta-evaluation questionnaire, containing five sections. The first four sections incorporate evaluation statements to assess the framework in terms of its evaluation of instructional design principles, VR design principles, usability and context-specific design aspects, respectively. The fifth section deals with the effectiveness of heuristic evaluation for evaluating interactive VR e-training systems.

4. **Evaluate in practice**:
   The meta-evaluation was conducted by administering the meta-evaluation questionnaire to the same six expert evaluators who had conducted the heuristic evaluations of the two prototype VR training programs, *LSF* and *ISGC*. They were selected as a purposive sample due to their experience in applying the evaluation framework.

5. **Reflection**:
   The researcher interpreted the findings of the meta-evaluation and revised the framework accordingly. This final DBR cycle was a demonstration of the versatility of the DBR process, in that the researcher not only evaluated the VR training products, but also refined the instrumentation and increased its sensitivity so as to strengthen future VR training artefacts.

A summary of the research methods applied in DBR Cycle 4 is presented in Table 5.7.

Table 5.7: Research Design: Cycle 4 research methods.

Cycle 4		**Outputs**
Participants	• 2 Usability Experts • 2 Mine Training Experts • 2 VR Development Experts	• Literature review on meta-evaluation • Meta-evaluation questionnaire • Improved *DEVREF* Framework
Data Collection Methods	Survey research	**Data**
Data Analysis	Quantitative – elementary statistical analysis Qualitative – tabulating data, organised by themes, calculating frequencies of occurrences and responses	Heuristic evaluation by 6 experts

5.6.8. Chronological timeline of the study

Figure 5.7 shows the activities and dates of the research in chronological order. The research process took several years to complete due to the iterative nature of DBR, the complexity of the real-world problem being researched, the application of various research methods, the novelty of the technological interventions, and the execution of four DBR cycles. The involvement of students in the development of the prototypes also contributed to lengthy development times.

The dates of these activities, as presented in Figure 5.7, serve to explain why most references cited in the description of the research activities and associated findings are relevant to particular time periods, although such descriptions were also updated with newer references, where applicable.

Figure 5.7 lists the main activities in the study. The writing up and improving of the thesis was done throughout and continuously.

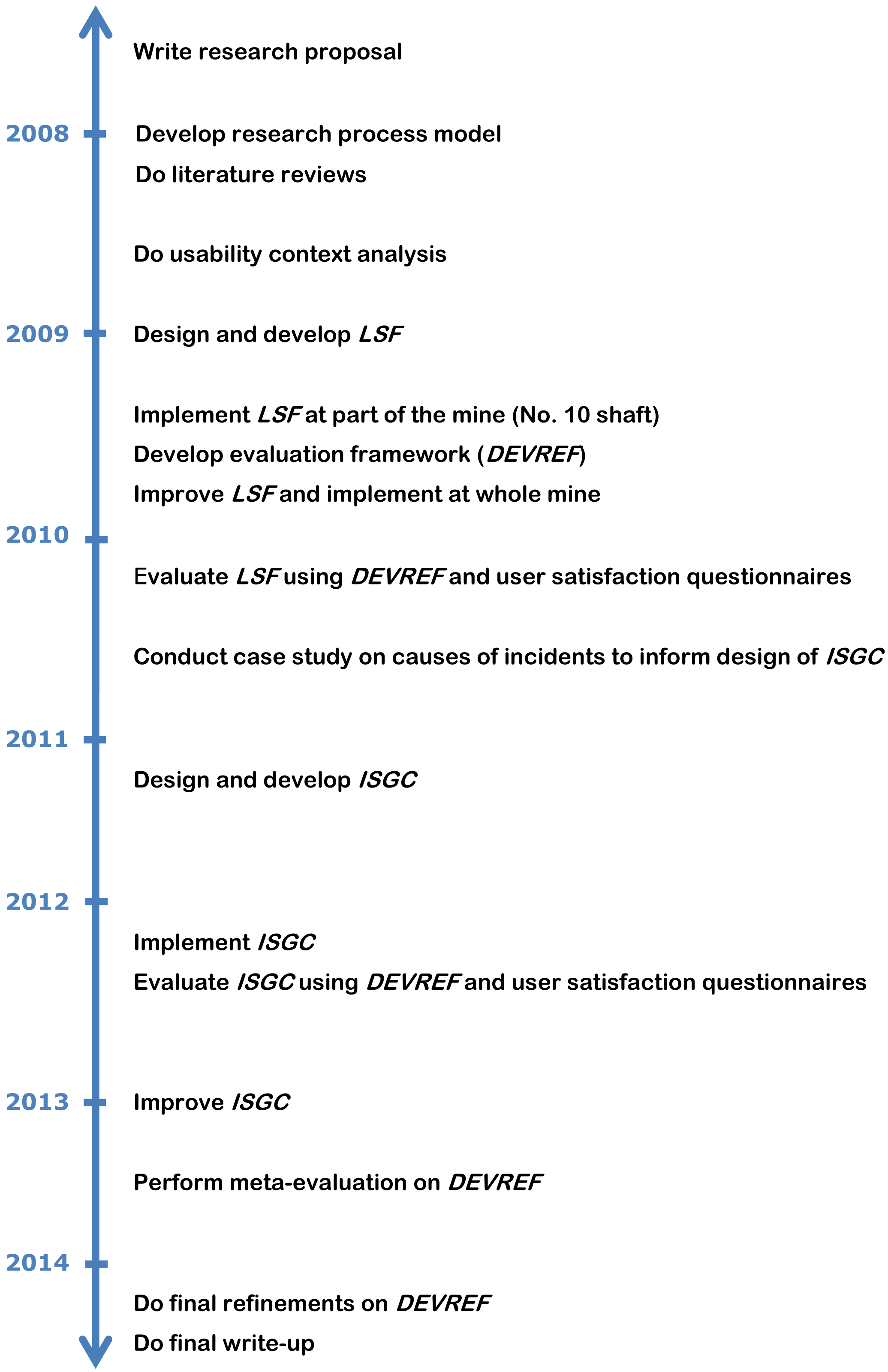


Figure 5.7: Chronological timeline of the study

5.7. Research methods

The research methods used in this study were: literature studies, prototyping, survey research, heuristic evaluation, case study research and informal participant observation. The main strategies were the methods employed for the evaluations of the prototypes, that is, heuristic evaluation and survey research, while the other methods were applied in supporting roles. Due to the study's focus on evaluation, more than one method was used to strengthen the process, an end-user survey and an expert evaluation method. The following subsections provide more detail on each of the research methods.

5.7.1. Literature Study

A literature study usually entails a review of the literature dealing with the chosen topics. This provides pertinent secondary data and sets the scene for a clear formulation of the research problem (Welman & Kruger, 2001). It brings the reader up to date on previous research and related work in the areas relevant to the study, and can also point out agreements and disagreements among previous researchers (Babbie, 2010). The literature review may also reveal inconsistencies and gaps that justify further research. Furthermore, it enables researchers to indicate exactly where the proposed research fits in.

Chapters Two, Three and Four are literature studies relating to the use of VR in training, relevant human computer interaction aspects, and training practices in the South African mining industry, respectively.

5.7.2. Prototyping

The term *prototype* refers to a simplified program or system that serves as a guide or example for the complete program or system (Olivier, 2004). Though programming *per se* is not research, prototyping can be applied to demonstrate that a new model or method can indeed be implemented. Prototypes serve as vehicles for experimentation and the construction of the prototype can also provide new insights. Associated research occurs when the prototypes are investigated and evaluated both in real-world use and while not being used by the target group.

Snyder (2003) and Dumas and Fox (2007) distinguish between paper prototypes, also called low-fidelity prototypes, and fully functional (high-fidelity) prototypes. Paper prototypes are low cost, involve little development time and multiple designs can be presented rapidly for evaluation, but it may be difficult to assess the actual look and feel of the interface as it does not allow for user interactions. Fully-functional prototypes require more development time and intricate programming, but allow realistic user interactions. Fully-functional prototypes are preferred in scenarios where presentations are made to management or clients to demonstrate actual products, or for collection of user performance data, for example, time to complete tasks or evaluating realism of graphics. A possible disadvantage of using fully-functional prototypes is that users may be hesitant to criticise, since they may feel that the system is already a finished product and their feedback may not be relevant (Lazar, Feng & Hochheiser, 2010).

According to Olivier (2004), prototypes can be useful in different ways.

- Proof of concept: A prototype can be constructed to demonstrate that a proposed concept actually works.
- Experimentation: A prototype can be used to determine facts about the prototyped system, for example, user reactions or system performance.
- Visualising abstract ideas: When constructing a model, prototyping can be used to focus the designer's mind on details that could easily have been overlooked.
- Exploratory research: Existing concepts can be demonstrated in new application areas.

For this research, prototyping was important to demonstrate implementation of the proposed training models and for experts and users to be able to evaluate the systems. The prototypes developed were used for proof of concept, experimentation and exploratory research.

5.7.3. Survey research

Survey research, also called query research (Dix, Finlay, Abowd & Beale, 2004), includes questionnaires and interviews. The use of questionnaires and interviews as a usability evaluation method was discussed in Section 3.6.1.3.

According to Babbie (2010:252), survey research is "probably the best method available to the researcher who is interested in collecting original data for describing a population too large to observe directly", and "surveys are also excellent vehicles for measuring attitudes and orientations in a large population". These characteristics make survey research ideal for this research among end-users in this study, where questionnaires and semi-structured interviews were used to collect data at several stages of the study, as indicated in Tables 5.4 to 5.7. Survey research was also applied in the form of heuristic evaluation, where assessment of the prototypes and the evaluation framework itself was performed by domain experts. More details on heuristic evaluation was discussed in Chapter Three, Section 3.6.2.3.

5.7.3.1. Questionnaire surveys

Surveys may be used for descriptive, explanatory and exploratory purposes. Within the field of Human computer Interaction, Lazar *et al.* (2010) list the following advantages of questionnaire surveys:

- It is easy to collect uniform data from a large number of people.
- Questionnaires can be used for many different research goals.
- They are useful in providing an overview of a population.
- When done on paper, they do not require special tools or equipment.
- Surveys are relatively unobtrusive.

Rogers, Sharp and Preece (2011) point out that it can be more difficult to develop good questionnaire questions than structured interview questions because with questionnaire surveys, the interviewer is usually not available to explain or to clarify any ambiguities. On the other hand, questionnaires have the advantage of reaching a wider subject group than interviews, and are inexpensive and easy to use (Dix *et al.*, 2004; Shneiderman & Plaisant, 2005).

Further drawbacks of using surveys include difficulty in getting detailed information, difficulty in following-up interesting trends that may emerge during analysis of collected data afterwards, and issues relating to the literacy levels of respondents when answering the survey questions (Babbie, 2010).

5.7.3.2. Interview surveys

An interview is a data-collection encounter in which one person (an interviewer) poses questions to another person (the interviewee or respondent), which can be done face-to-face or by telephone (Babbie, 2010). Cohen, Manion and Morrison (2011) describe interviews as a flexible and powerful tool for data collection. According to Oates (2006), interviews as a data generation method are particularly suitable when a researcher wishes to obtain detailed information, ask complex or open-ended questions, or explore aspects not easily observed or described via pre-defined questionnaires. Types of interviews include the following

- Unstructured or informal interviews: Interview questions are not pre-planned and emerge from the immediate context. The interviewer introduces a topic and then the interviewees are allowed to develop and express their own ideas and detailed responses.
- Semi-structured or guided approach: Topics and issues to be covered are specified in advance, but the order of questions may change depending on the flow of the interview and whether issues are raised for which questions were not preprared.
- Structured or closed interviews: The wording and sequence of questions are determined in advance and are repeated in the same order for all interviewees.

(Cohen *et al.*, 2011; Lazar, Feng & Hochheiser, 2010; Oates, 2006).

Interviews can be exploratory and flexible, but this method does have its challenges. The following are possible drawbacks of interviews:

- They are expensive in terms of time taken.
- They are open to interviewer bias.
- They are more difficult to conduct than surveys and require the interviewer to have interviewing skills, which can take significant practice to develop.
- They may be inconvenient for respondents.
- Interviewee fatigue may hamper the interview.
- Anonymity may be difficult to ensure.

(Cohen *et al.*, 2011; Lazar *et al.*, 2010; Rogers *et al.*, 2011).

5.7.4. Heuristic evaluation

As discussed in Section 3.6.2.3, heuristic evaluation is an inspection technique whereby a small number of experts apply a set of usability principles called *heuristics*, to evaluate whether a user interface conforms to these principles (Hix & Gabbard, 2002; Lazar *et al.*, 2010; Madan & Dubey, 2012). According to Rogers *et al*. (2011), the way in which experts are intended to use these heuristics is by judging them against aspects of the interface. Heuristic evaluation is appropriate for use in various development phases, from initial prototype to early design evaluations (Nielsen, 2005; Shneiderman and Plaisant, 2005; Ssemugabi and de Villiers, 2010). Heuristic evaluation is effective in identifying problems in systems (Zaibon & Shiratuddin, 2010).

According to Karoulis and Pombortsis (2003), heuristics are normally derived from academic and professional research studies, existing criteria lists, or field observations and prior experience in the given domain. In addition to these general heuristics that should be considered for all interfaces, the evaluator can add customised usability heuristics pertaining to the specific domain for which the application is developed (Rogers *et al.*, 2011).

This study proposes an evaluation framework for the heuristic evaluation of desktop VR training systems for the mining industry. The heuristics of the framework are derived from academic literature and the personal experience of the present researcher, and include heuristics specific to safety training in the mining industry. These heuristics are presented in Section 5.8. In this study, heuristic evaluation was applied to evaluate two VR prototype systems, as well as to perform a meta-evaluation on the proposed evaluation framework. These evaluations were performed by a set of six experts: two usability experts, two mining training experts and two VR development experts, with varying expertise. The two usability experts had both completed Master's degrees in usability and were enrolled for PhD studies. Both of them had previous experience in usability evaluation. The mining training experts had many years of mine training experience and both were heads of their respective departments. The VR developers had been involved in such developments for the previous seven years.

5.7.5. Case study research

A case study is defined as an investigation to answer specific research questions which seeks a range of different evidence from the case settings (Gillham, 2000). This evidence can be abstracted and collated to obtain the best possible answers to the research question (Yin, 2002).

A case study investigates a contemporary phenomenon within a real-life context and can provide qualitative and/or quantitative data (Olivier, 2004). A case study is often used in such a situation to explain causal links in real-life situations when it is difficult, complex or impossible to use other research methods such as experiments (Gillham, 2000). In such cases, the data obtained would be more comprehensive than that obtained from a survey among a sample of the population. In the case study described in Chapter Seven, details relating to falls of ground at one of the world's largest platinum mines were analysed and discussed.

5.7.6. Informal participant observation

Participant observation requires the researcher to take part in, and report on, the experiences of the members of a group, community or organisation involved in a process or event (Welman & Kruger, 2001). The participant observer becomes a member of the group or event being studied, in order to personally experience what the group members experience, understand their environment, and comprehend the meaning and significance of their behaviour. The researcher therefore performs a dual role of experiencing the activities of the group and also observing and recording such experiences.

Rogers *et al.* (2011) differentiate between degrees of observer participation, ranging from *insider* at one end of the spectrum and *outsider* at the other. A participant observer at the insider end of the spectrum attempts to become a full member of the group being studied, whereas an observer on the outsider end of the spectrum does not take part in the study environment and is called a passive observer. Similarly, Cresswell (2009) distinguishes between researcher roles in observation.

- Complete participant: The researcher conceals his/her role and participates in the activities being observed.
- Observer as participant: The researcher participates in activities but the role of the researcher as observer is known.
- Complete observer: The researcher observes without participating.

In this study, informal participant observation was done by the researcher as complete observer for the usability context analysis described in Chapter Six, as well as observing trainees using the *LSF* prototype at No. 10 shaft, as mentioned in Section 5.6.7.

5.7.7. Application of the research methods in this study

Figure 5.8 is an extended version of Figure 1.3 in the introductory chapter and indicates in green where the various research methods were used in this study.

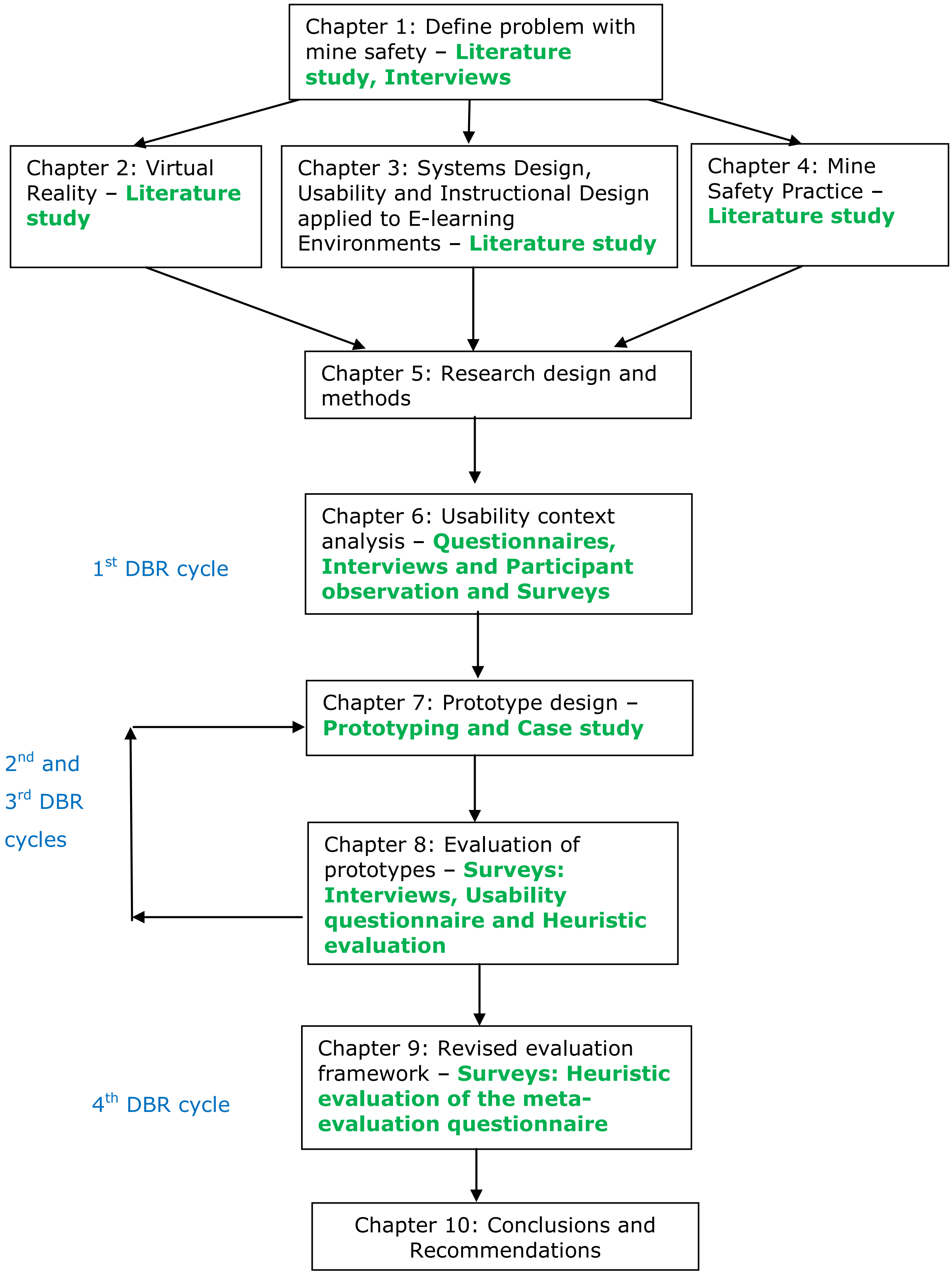


Figure 5.8: Application of Research Methods.

5.8. Synthesised evaluation framework for heuristic evaluation of desktop VR training applications

An evaluation study is founded on an evaluation method or methods (see Section 5.7) and on a set of evaluation criteria. As mentioned in Section 3.6.3, this study synthesises a new set of criteria specifically customised for evaluation of desktop VR training applications within the context of mining safety. Hence the framework is called the Desktop Virtual Reality Evaluation Framework (*DEVREF*). Since one of the major evaluation methods in this study is heuristic evaluation, the criteria are termed 'heuristics'.

This section addresses Research Subquestion 5 of this study, as does Section 3.6.3. The question is answered in detail in Section 9.4.

RQ5	What structure, categories and criteria should be incorporated in an evaluation framework for virtual reality training systems in the mining industry?

Since this thesis relates to the development of a framework for evaluating VR training systems systems, this section briefly addresses the concept of frameworks in general and the nature of the framework developed for this research, before describing the evaluation framework itself. First, the distinction is drawn between conceptual frameworks and evaluation frameworks.

A *conceptual framework* is often depicted graphically, representing a set of objects under under investigation, namely, the constructs and variables, and the relationships between them (Miles & Huberman, 1994). This relates to studies where the problem cannot meaningfully be researched in reference to only one theory, or concepts resident within one theory. In such cases, the researcher may have to integrate existing views into a model or conceptual framework (Imenda, 2014). Leshem and Trafford (2007) consider the role of conceptual frameworks in the context of conceptualisation in doctoral research. They similarly explain that the term, conceptual framework, is mainly used to describe a particular function and a set of interrelationships in a research process.

An *evaluation framework* is a simpler structure, usually presenting categories of evaluation criteria in a tabular format (Ssemugabi & De Villiers, 2010; Vavoula & Sharples, 2009). It is less focused on complex interrelationships, and demonstrates an understanding of theories and concepts that are relevant to the topic of the research and that relate to the broader areas of knowledge being considered (Labaree, 2015).

This research introduces a single, integrated evaluation framework, developed for the purpose of evaluating interactive desktop VR training systems. The work commenced with a study of existing frameworks and models, as discussed in Section 1.5.3. Current evaluation frameworks are limited, because they are either confined to evaluation of a specific type of virtual environment or they focus on a restricted aspect of virtual environments. This study addresses the gap for a framework for evaluation of desktop VR training systems for the mining industry, by investigating the design and development of such systems meticulously and comprehensively from the following perspectives: instructional design; usability; and VR systems design, situated in the context of underground mining. These different perspectives are integrated into a single framework, providing a multi-faceted evaluation approach.

The proposed evaluation framework consists of four categories of heuristics, which can also be termed performance criteria (see Section 5.6.5).

- Category 1: Instructional design – includes heuristics related to pedagogical effectiveness, learning theories and multimedia learning design.
- Category 2: General usability – includes interface design and interaction, and heuristics that support the goals of usability.
- Category 3: Virtual reality system design – includes heuristics specific to the design of virtual reality systems.
- Category 4: Context-related heuristics – includes heuristics related to the content and the application domain.

The framework is derived from the literature, as described in Chapter Three, as well as from the personal experience of the researcher, who has been involved for the past ten years in the design and development of virtual reality training systems for the mining industry. Teräs and Herrington (2014) warn against the common pitfall of simply adapting new technology to traditional systems, practices, and methods, rather than using authentic learning principles that complement the affordances and characteristics of the technology. This statement strengthens the case for the inclusion of a category of heuristics relevant to the design of virtual reality training systems.

The structure of the evaluation framework, the categories, the heuristics and the associated literature references are set out in Table 5.8. Some of the authors listed in the third column are cited in the text (Chapters Two to Five); others are mentioned only in the table. As indicated in the chronological timeline of the study (Figure 5.7), the *DEVREF* Framework was developed by the researcher during 2009 and 2010, which accounts for the dates of the references in Table 5.8. The various studies in this research led to the improvement of *DEVREF* and a revised version is presented in Table 9.6 in Chapter Nine. The revised framework also cites more recent sources, which confirm and extend the original framework.

Table 5.8: Heuristic Evaluation Framework for Desktop VR Training Applications.

	Category 1: Instructional Design	
	Heuristic/Criterion	**Literature References**
1	**Clear goals, objectives or outcomes:** • The training program makes it clear to the learner what is to be accomplished and what will be gained from its use. • There are clear goals, objectives or outcomes for the training program. • Clear goals, objectives or outcomes are communicated at the beginning of the training program. • The outcomes are measurable.	Ritchie and Hoffman (1997), Albion (1999), Wein, Piccirilli, Coffey and Flemming (2000), Alessi and Trollip (2001), Reeves, Benson, Elliot, Grant, Holschuh, Kim, Kim, Lauber and Loh (2002), McLoughlin, in Edmundson (2003), Ardito, Costabile, De Marsico, Lanzilotti, Levialdi, Plantamura, Roselli, Rossano and Tersigni (2004b).
2	**Instructional assessment:** • The program provides assessment opportunities that are aligned with the objectives or outcomes. • The assessment opportunities will serve to enhance trainees' performance.	Albion (1999), Patel, Stefani, Sharples, Hoffmann, Karaseitanidi and Amditis (2006).
3	**Feedback to user responses:** • The training program provides trainees with constructive and supportive feedback on their performance. • The feedback is relevant to the training content. • The feedback informs the trainee regarding his level of achievement in the training program. • The feedback indicates incorrect responses and provides information on the correct responses.	Alessi and Trollip (2001), Vrasidas (2004).

4	**Motivation and creativity**: • The system supports intrinsic motivation by providing challenges to trainees and encouragement when errors are made. • The program captures the trainee's attention early and retains it throughout. • This training program increases trainees' confidence by providing them with reasonable opportunities to accomplish the objectives successfully. • The program engages trainees by its relevant content. • The program engages trainees by its interactivity.	Albion (1999), Alessi and Trollip (2001), Reeves *et al.* (2002), Chalmers (2003), Vrasidas (2004), De Villiers (2005a), Ssemugabi and De Villiers (2010).
5	**Differences between individual users** • The system takes account of linguistic and cultural differences by allowing trainees to select between different languages. • The system caters for trainees with different levels of expertise regarding the content. • The system caters for trainees with different levels of computer experience.	Alessi and Trollip (2001), Barber (2002), Reeves *et al.* (2002), Chalmers (2003), Liu, in Edmundson (2003), McLoughlin, in Edmundson (2003), Kamppuri, Tedre and Tukiainen (2006).
6	**Reduction of extraneous processing in working memory:** • The training program effectively uses signalling to highlight essential issues, such as restating important points, using headings for important points, or stressing them in audio mode. • Redundancy is avoided by not presenting unnecessary information. • Redundancy and overload are avoided by not reiterating the same material in multiple modes (e.g. the program presents information using pictures and spoken words, rather than presenting it in pictures, spoken words, and printed words).	Chalmers (2003), Mayer (2008), Hollender, Hofmann, Deneke and Schmitz (2010).
7	**Fostering of germane cognitive load (germane cognitive load is the load devoted to the processing, construction and automation of mental schemas):** • The training program supports the formation of mental schemas by explaining where newly acquired knowledge fits into the bigger picture. • The system encourages encoding of the training content into long-term memory by presenting	Albion (1999), Sweller (1999), Alessi and Trollip (2001), Chalmers (2003), Ardito *et al.* (2004b),

	questions after each learning segment. • Sufficient scaffolding support is provided (in the form of hints, prompts and feedback) to help trainees achieve training goals. • The training program presents narration in a colloquial conversational style. • The training program prompts trainees to link concrete example information to more abstract information.	Van Merriënboer and Sweller (2005), Sawicka (2008), Bennett, Stothard and Kehoe (2010), Hollender *et al.* (2010).
8	**Appropriate intrinsic cognitive load**: • Working through the training program does not cause trainees to divide their attention between multiple sources of visual information. • The program enhances retention by presenting information in learner-paced segments, rather than as a continuous presentation. • The system effectively supports dual-channel processing of simultaneous visual and verbal material.	Alessi and Trollip (2001), Pollock, Chandler and Sweller (2002), Reeves *et al.* (2002), Chalmers (2003), Shneiderman and Plaisant (2005), Mayer (2008), Zhang, Wang, Zhao, Li and Lou (2008).
Category 2: General Usability		
	Heuristic/Criterion	**References**
1	**Functionality**: • The interface provides the level of functionality the user requires to complete a task. • The interface provides adequate functionality to return to a previous screen. • Icons, labels and symbols are intuitive and meaningful to trainees, bearing in mind the level of trainee context and experience.	Kalawsky (1999), Dringus and Cohen (2005).
2	**User guidance:** • The interface provides clear indications of what the next required action will be. • Help for operating the program is accessible at any time and appropriate. • Trainees receive clear instructions on how to use the training program. • Guidance to solve problems is given visually as examples, diagrams, videos or photographs.	Nielsen (1994), Kalawsky (1999), Alessi and Trollip (2001), Reeves *et al.* (2002), Dringus and Cohen (2005).
3	**Consistency:** • There is consistency in the sequence of actions taken in similar situations.	Nielsen (1994), Kalawsky (1999),

	• There is consistency in the use of images, prompts, screens, menus, colours, fonts and layouts. • Objects, options, and permissible actions are visible so that users do not have to remember instructions. • Different screens that have similar operations, use similar elements for achieving similar tasks.	Squires and Preece (1999), Reeves *et al.* (2002), Dix *et al.* (2004), Dringus and Cohen (2005), Shneiderman and Plaisant (2005), Wong, Marcus, Ayres, Smith, Cooper and Paas (2009).
4	**Error correction:** • Error messages are expressed in plain language. • Learners are provided with the necessary help to recover from cognitive errors. • Error messages indicate precisely what the problem is and give simple, constructive, specific instructions for recovery.	Nielsen (1994), Kalawsky (1999), Powell (2001), Reeves *et al.* (2002), Karoulis and Pombortsis (2003), Dix *et al.* (2004), Shneiderman and Plaisant (2005).
5	**System status**: • The program keeps the trainee informed about what is going on through constructive, appropriate and timely feedback. • For every action taken by the trainee, there is a visual or audio response by the training program, so that learners can see and understand the results of their actions. • The program responds to actions initiated by the user and there are no surprise actions from the system's side.	Nielsen (1994), Levi and Conrad (1996), Albion (1999), Squires and Preece (1999), Reeves *et al.* (2002), Dix *et al.* (2004), Shneiderman and Plaisant (2005).
6	**Error prevention:** • The training program is designed in a way that the learner cannot easily make serious errors. • When the learner makes an error, the system responds with an error message, to prevent further similar errors. • Trainees can recognise situations where errors are due to the way they provided input, and not due to incorrect content in their response. • The system is robust and reliable throughout.	Nielsen (1994), Squires and Preece (1999), Alessi and Trollip (2001), Powell (2001), Reeves *et al.* (2002), Karoulis and Pombortsis (2003), Dix *et al.* (2004), Dringus and Cohen (2005), Shneiderman and Plaisant (2005).
7	**Aesthetics:** • The screens are pleasing to look at. • The buttons and selections are of an adequately viewable size. • The text is of an adequately viewable size. • There is not too much content or information on the screens.	Alessi and Trollip (2001), Reeves *et al.* (2002), Dringus and Cohen (2005).

8	**Interactivity:** • The training program uses clear and simple terminology that supports trainees in understanding how to interact with the system. • The interactions provided by the program support trainees in learning the necessary content. • Working through the program requires regular trainee interactivity to maintain attention and facilitate comprehension.	Alessi and Trollip (2001), Preece, Rogers and Sharp (2002), Dringus and Cohen (2005).
	Category 3: Virtual Reality System Design	
	Heuristic/Criterion	**References**
1	**User control:** • The user is able to interact with, or control, the virtual environment in a natural manner. • Responses from the environment to the participant's control actions and movements, are perceived as immediate or close-to-immediate. • The system permits easy reversal of actions. • Trainees are able to exit the system at any time when they need to do so.	Nielsen (1994), Kalawsky (1999), Squires and Preece (1999), Dix *et al.* (2004), Shneiderman and Plaisant (2005), Wilson and D'Cruz (2006).
2	**Multimodal system output/feedback:** • The effect of the trainee's actions on objects in the virtual environment, is immediately visible and conforms to the laws of physics and the trainee's perceptual expectations. • The visual representation of the virtual world maps to the trainee's perception of that environment. • There are no major distortions in visual images. • Audio is integrated seamlessly into user task activity. • Audio information is meaningful and timely.	Mereu and Kazman (1996), Oshhima, Yamamoto and Tamura (1996), Richard, Birebent, Coiffet, Burdea, Gomex and Langrana (1996), Kalawsky (1999), Hix and Gabbard (2002), Sutcliffe and Gault (2004).
3	**Presence:** • Users feel as if they are part of the virtual environment and not isolated from it. • The virtual environment experience is consistent with similar real-world experiences.	Witmer and Singer (1998), Kalawsky (1999), Sadowski and Stanney (2002), Bowman and McMahan (2007).

4	**Orientation:** • Users do not find it difficult to maintain awareness of their location while moving through the virtual environment. • The virtual environment includes appropriate spatial labels and landmarks to support user orientation. • It is clear to the user how to exit the virtual environment.	Darken and Sibert (1996a, 1996b), Marsh, Wright and Smith (2001), Stanney, Mollaghasemia, Reevesa, Breaux and Graeber (2003), Sutcliffe and Gault (2004), Bennett *et al.* (2010).
5	**Navigation:** • It is easy for users to move and reposition themselves in the virtual environment. • Means of navigation are consistent throughout the system.	Squires and Preece (1999), Bowman, Kruijff, LaViola and Poupyrev (2001), Stanney *et al.* (2003), Kalawsky (1999), Alessi and Trollip (2001).
6	**Object interaction – selection and manipulation**: • Input devices are easy to use and easy to control. • Object interactions are designed realistically to reproduce real-world interaction. • The system provides the ability to rotate 3D objects and increase the level of detail when necessary for task performance.	Witmer and Singer (1998), Kalawsky (1999), Bowman *et al.* (2001), Stanney *et al.* (2003).
7	**Fidelity:** • The simulations in the system are accurate. • The objects in the virtual environment move in a natural manner. • The virtual environment displays adequate levels of realism. • High-fidelity graphics are used where required.	Kalawsky (1999), Sutcliffe and Gault (2004), Bennett *et al.* (2010).
8	**Variety in user modes:** • The system employs various modes to cater for a range of users from novices to experts. • The system provides various user-guidance modes, e.g. Free mode, Presentation mode, Guided mode and Discovery mode.	Arendarski, Termath and Mecking (2008), Bennett *et al.* (2010).

Category 4: Context-specific criteria		
	Heuristic/Criterion	**References**
1	**Authentic tasks:** • The training system supports particular work practices in the context of their natural work environment. • The system is customised according to the curriculum and learner-specific needs. • The program includes tasks applicable to the job context of the trainee.	Sachs (1995), Beyer and Holtzblatt (1998), Harris and Henderson (1999), Jonassen (1999), Squires and Preece (1999), Notess (2001), Reeves *et al.* (2002), Edmundson (2003), Ardito *et al.* (2004b), Chen, Toh and Fauzy (2004), Vrasidas (2004), Ssemugabi and De Villiers (2010).
2	**Appropriate reference materials:** • The system includes supplementary reference materials, providing information to trainees on standard operating procedures used in the application domain. • The reference materials included in the system are relevant to the problem scenarios. • The reference materials are at a level appropriate to the trainees.	Albion (1999), Alessi and Trollip (2001).
3	**Comprehensive scope:** • The learning material covers all the vital aspects relating to the topics being addressed. • The training covers possible consequences of trainees not applying the learning material correctly in their work place.	Experience of the present researcher
4	**Adaptive design:** • The design of the training system is adaptive to changes in site practices. • The system refers to the current standard operating procedures. • The system randomises assessment details such as questions and multiple-choice answers when administering assessment.	Experience of the present researcher

5	**Relevant subject matter:** • The subject matter matches the goals and objectives of the training program. • The subject matter is presented in an appropriate structure. • The information provided in the program is accurate. • The system 'speaks the trainee's language' by using terms, phrases, symbols and concepts familiar to the trainee and common to the application domain. • The level of language use, in terms of grammar and style, is appropriate for the target audience.	Nielsen (1994), Squires and Preece (1999), Alessi and Trollip (2001).
6	**Trainee preparedness:** • Trainees are shown how to use the software prior to undergoing training with the program. • PC literacy pre-training is available to trainees who are not comfortable with using computers for training.	Hollender *et al.* (2010)
7	**Appropriate record keeping:** • The system maintains trainee records and assessment results. • The system monitors and displays trainee progress. • The system ensures legal compliance by capturing detailed individual performance data.	Vrasidas (2004)
8	**Understandable and meaningful symbolic representation:** • Symbols, icons and terminology that represent concepts and objects are used consistently throughout the program. • Symbols, icons and terminology are intuitive within the context of the task. • Metaphors correspond to real-world objects or concepts.	Nielsen (1994), Squires and Preece (1999), Alessi and Trollip (2001), Stanney *et al.* (2003), Dix *et al.* (2004), Oviatt (2006).

The *DEVREF* Framework presented in the comprehensive preceding table, was used for the heuristic evaluation of the *LSF* and *ISGC* prototypes, as mentioned in the discussion of DBR Cycles 2 and 3 in Section 5.6.7. The design of the evaluation instrument and the findings of these evaluations are presented in Chapter Eight.

5.9. Validity, reliability and triangulation

This section defines the theoretical concepts of validity, reliability and triangulation. The concepts are revisited in Section 10.5 in Chapter Ten, where it is outlined how each of the three concepts is implemented in this research.

5.9.1. Validity

Validity is described as the degree to which a research study measures that which it was intended to measure. For qualitative data, validity can be addressed through the honesty, depth, richness and scope of the data, the participants involved, the extent of triangulation and objectivity of the researcher (Winter, 2000). For quantitative data, validity can be improved through careful sampling, appropriate data collection instruments and appropriate statistical analysis of the data. The findings must accurately describe the phenomena being researched (Cohen *et al.*, 2011).

Both quantitative and qualitative methods can address validity. Within qualitative methods, Maxwell, as cited by Cohen *et al*. (2011), describes five types of validity.

- Descriptive validity refers to the factual accuracy of the data.
- Interpretive validity gives the meaning and interpretation of data according to participants.
- Theoretical validity refers to the explanations of phenomena by the researcher.
- Generalisability involves understanding the usefulness of the research in other situations.
- Evaluative validity refers to an evaluative or judgemental approach by the researcher.

Several different kinds of validity are described in the literature. The following subsections briefly explain some of the frequently-mentioned types which are relevant to this study: Internal validity, External validity, Content validity, and Construct validity.

5.9.1.1. Internal validity

Internal validity reflects the extent to which a causal conclusion based on a study is warranted. This means that an explanation of an event, issue or set of data can actually be sustained by the data and that the findings should accurately describe the phenomena being researched. Gliner & Morgan (2000), Oates (2006) and Cresswell (2009) identified several threats to internal validity.

1. History: Environmental events outside of the study may influence participants' responses.
2. Maturation: Participants or subjects may change during the course of the study or even between measurements.
3. Testing: Repeatedly measuring the participants may improve their performance due to the repeated exposure and not necessarily due to the intervention. Participants may remember the correct answers or may be conditioned to know that they are being tested.
4. Instrumentation: In conducting experiments, the calibration of instruments can change between experiments. This also refers to different human evaluators giving different results, or observers having unconsciously changed the criteria they use to make judgements.
5. Statistical regression: This may occur when participants are selected on the basis of extreme scores. Such participants will tend to have less extreme scores if re-tested and would likely evolve into a more normal distribution with repeated testing.
6. Differential selection: If the researcher allows participants to pick their own group, the two groups may differ in some way that influences the outcomes.
7. Experimental mortality: This may occur if inferences are made on the basis of only those participants that have participated from the start to the end. However, participants may have dropped out of the study before completion, and possibly even due to the study or experiment itself.
8. Selection-maturation interaction: This occurs if biases in assignment interact differentially with maturation or other factors.
9. Diffusion of treatment: Participants in the control and experimental groups may influence each other if they are allowed to communicate.
10. Compensatory demoralisation and rivalry: When only the experimental group receives treatment (e.g. therapy) and the control group receives nothing, then the benefits of the experiment may be unequal or resented, or the participants in the control group can feel devalued.

5.9.1.2. External validity

External validity refers to the degree to which the results obtained from a small sample group can be generalised to the wider population. To prevent poor external validity, any research design must justify sampling and selection methods (Shuttleworth, 2009). Some threats to external validity include the following:

- Selection effects: Constructs selected are only relevant in a certain group.
- Setting effects: Situational specifics factors such as time, location, scope and extent of measurement may potentially limit generalisability.
- Reactivity effects: Results might not be generalisable to other settings or situations if the effects found only occurred due to studying the situation, also known as Hawthorne effects.
- History effects: Results are due to unique circumstances.

5.9.1.3. Content validity

Content validity refers to how much a measure covers the range of meanings included in a concept. A research instrument demonstrating content validity should fairly and comprehensively cover the domain it claims to cover. In the case of each issue in a particular domain not being addressed in its entirety, the researcher must ensure that the elements of the main issue to be covered in the research are a fair representation of the wider issue under investigation. Furthermore, the elements chosen for the research sample should then be addressed in detail (Babbie, 2010; Cohen et al., 2011). Cresswell (2009) indicates that pilot testing of a survey instrument could assist to establish content validity.

As is the case in this study, as an example, if user satisfaction is evaluated by a questionnaire then the user satisfaction questionnaire should cover all the aspects of user satisfaction. This implies that the researcher should be familiar with the research domain before designing the research instrument (Oates, 2006). This is usually accomplished by studying available literature and/or having practical experience in the research domain.

5.9.1.4. Construct validity

Where content validity is concerned with whether the questions in a survey instrument are a well-balanced sample of the domain under investigation, construct validity refers to whether the instrument is actually measuring the characteristic being investigated. This means that the articulation of the survey questons is important (Cohen *et al.*, 2011). To ensure construct validity, Oates (2006) states that survey questions should be brief, relevant, unambiguous, specific and objective. Threats to construct validity occur when researchers use inadequate definitions and measures of variables (Cresswel, 2009).

5.9.2. Reliability

Whereas validity refers to getting results that accurately reflect the concept being measured, reliability means getting consistent results from the same measure. It is important that reliability and validity are addressed at all stages of a research study, specifically during design and methodology, sampling, timing, data collection, data analysis and data reporting.

Reliability refers to the dependability and consistency of the research findings. When research is reliable, a similar group of respondents in a similar context would produce similar results (Lazar *et al.*, 2010), and a particular technique, applied repeatedly to the same object, would yield the same result each time (Babbie, 2010). According to Oates (2006), questions in a reliable survey instrument should be neutral and not lead respondents into answering a certain way.

According to Cohen *et al.* (2011), reliability in quantitative research often concerns consistency, accuracy, predictability, equivalence, replicability, concurrence, descriptive and causal potential. On the other hand, reliability in qualitative research often concerns accuracy, fairness, dependability, comprehensiveness, respondent validation, empathy, uniqueness, explanatory and descriptive potential, and confirmability. In all research studies, reliability can be improved by:

- minimising external sources of variation;
- standardising conditions under which measurement occurs;
- improving researcher consistency;
- broadening the set of measurement questions by including similar questions within the instrument, by increasing the number of researchers (triangulation), and by increasing the number of measurement occasions; and
- excluding extreme responses (termed outliers).

5.9.3. Triangulation

Triangulation refers to the use of more than one approach to the investigation of a research question, or the use of more than one data generation method, in order to enhance confidence in the ensuing findings (Oates, 2006). Data triangulation involves using different sources of information, which enables the researcher to use qualitative and quantitative data to corroborate each other. Such results can then be compared to improve reliability and validity (Cresswell, 2009). Triangulation within methods concerns the replication of a study as an indicator of reliability, whereas triangulation between methods involves the use of more than one method to improve validity.

The use of triangulation increases confidence in research data and may reveal unique findings and provide a clearer understanding of the problem due to the diversity and quality of data obtained. For example, using interviews as well as questionnaires adds a depth to results that would not have been possible using a single-strategy study, thereby increasing the validity and utility of the findings (Guion, Diehl & McDonald, 2011).

Cohen *et al.* (2011) and Oates (2006) distinguish between different types of triangulation.

- Time triangulation: Researchers collect data about a phenomenon at two or more different points in time.
- Space triangulation: Data is collected at two or more sites, e.g. different countries or cultures, to allow wider generalisation of results of studies.
- Combined levels of triangulation: More than one level of analysis is used, where levels are defined as individuals, groups and collectives (organisational, cultural or societal).
- Theoretical triangulation: Multiple perspectives are used to interpret a single set of data, which may involve the use of professionals outside of a particular field of study, even people from different disciplines.
- Investigator triangulation: More than one observer or investigator is involved in the analysis process, and the findings from each evaluator can then be compared to develop a broader and deeper understanding of how the different investigators view the issue.
- Methodological triangulation: Either the same method is used on different occasions, or different methods are used on the same object of study.
- Strategy triangulation: More than one research strategy is used in the same study.

Guion *et al.* (2011) also define environmental triangulation. This involves the use of different locations, settings, and other key factors related to the environment in which the study took place to identify which environmental factors, if any, might influence the information that is received during the study. These environmental factors are changed to see if the findings are the same across settings. If the findings remain the same under varying environmental conditions, then validity has been established.

The application of validity and reliability to the research, as well as the implementation of triangulation, are discussed in Chapter Ten.

5.10. Ethics

Ethical considerations come into play at three stages of a research project, namely:

- when participants are recruited;
- during the intervention or measurement procedure to which they are subjected; and
- in the release of the results obtained (Welman & Kruger, 2001).

Research ethics require that participants must be treated fairly and with respect. This means that they must be provided with information about the nature of the study which they can use to decide whether they want to be involved. Participation should be entirely voluntary and free from any implied or implicit coercion. Participants should sign an informed consent form, acknowledging that they are aware that they are taking part in a research project and giving their consent. Participants should also be assured that their privacy will be protected. Researchers should obtain consent for the collection and storage of personal information, identify the uses that will be made of any information, securely protect any information and limit the use and disclosure of such information. (Lazar *et al.*, 2010).

Participants should not be harmed in any way, regardless of whether they volunteer for the study. Apart from possible physical harm, there should be no revealing of information that could embarrass participants or endanger their careers, relationships or personal lives. Researchers can use anonymity or confidentiality to protect participants. A research project guarantees anonymity when the researcher cannot associate a given response with a specific respondent. This is difficult to achieve in interview surveys, as the interviewer collects the information from an identifiable respondent. When using questionnaires, assuring anonymity makes it difficult to keep track of who has or has not returned the questionnaires. If a research project guarantees confidentiality, then the researcher can identify a given person's response, but must not do so publicly. When a research project is confidential rather than anonymous, it is the researcher's responsibility to make that fact clear to the participants (Babbie, 2010).

In addition to their ethical obligations to participants, researchers also have ethical obligations to the scientific community concerning the analysis of the data and the way the results are reported. The researcher should be aware of a study's limitations and failures, and should make these clear to their readers. For example, negative findings should also be reported if they are related to the analysis, and not only strong, causal relationships among variables. Researchers should conduct research rigorously and with the correct procedures, report procedures and findings accurately and publicly, and, where applicable, avoid interference with the research by sponsors or those who give permission for the research to be undertaken (Cohen *et al*., 2011).

For this research, the researcher applied for ethical clearance from the Ethical Clearance Subcommittee of Unisa's College of Science, Engineering and Technology. This application included details regarding the location, objectives, research questions, research methods and the actual research instruments to be used. The research instruments also included consent forms to be signed by participants. Furthermore, the researcher undertook to carry out the study in strict accordance with the approved research proposal and the ethics policy of Unisa. The ethical clearance letter of approval is provided as Appendix A-1.

The researcher also obtained permission from the mine to conduct the research at the mine (Appendix A-2). A clear explanation of the research purpose and procedure was provided to participants prior to evaluations. Participants were asked to sign informed consent. As revealing their survey responses would not injure them in any way, it was decided not to use anonymity but rather confidentiality. Participants were ensured that, even though the findings of the evaluation would be used for research purposes and that the findings might be published in academic publications, their privacy would be protected by non-disclosure of their names, positions or affiliations. The informed consent document is given in Appendix B-2 as part of the user satisfaction questionnaire document.

For the heuristic evaluation, the six expert evaluators were each requested to sign a consent form. In the document, the evaluators acknowledged that their participation was voluntary, that they were assured of anonymity, and that their inputs would be used purely for academic reasons. This informed consent form was part of the heuristic evaluation instrument, which is attached as Appendix B-1.

5.11. Conclusion

The main purpose of this chapter was to provide a clear focus on the research design of this study with particular reference to the paradigm and methods used to answer the research questions, which are listed in Section 5.2. Design-based research was defined as an appropriate research paradigm for this study. Section 5.3 introduced design science and design research. Section 5.4 discussed design science research, while Section 5.5 dealt with the evolution of DBR from design experiments, design science, and development research. The characteristics of DBR were described and a consolidated summary of DBR features was provided.

The researcher presented a diagrammatic synthesised model for DBR and demonstrated how it was applied within the process flow of this study, involving the iterative design, development, evaluation and refinement of prototype virtual reality systems for e-training in the mining industry. The research design of the study (Section 5.6) indicated details of the application of four cycles of the DBR model, and the research methods used in the four cycles (Section 5.7). The *DEVREF* heuristic evaluation framework for evaluation of desktop VR training systems was presented in Section 5.8. This framework was applied in the evaluation of two prototype systems. The details of these evaluations are discussed in Chapter Eight. Section 5.9 explained how validity, reliability and triangulation were addressed in the study, while Section 5.10 considered the ethical issues.

This chapter addressed Research Subquestion 2: “Which research paradigm is most appropriate for the intended research?” in Sections 5.5 and 5.6. It also answered Research Subquestion 5: “What structure, categories and criteria should be incorporated in an evaluation framework for virtual reality training systems in the mining industry?” in Section 5.8. The deliverable of this chapter is therefore the synthesised DBR model.

The next four chapters give detailed descriptions of the activities of the four DBR cycles, as briefly outlined in Section 5.6.6. A colour-coded diagram at the beginning of each of these chapters, based on Figure 5.6 in this chapter, indicates which parts of which cycle are addressed by the chapter.

Chapter Six
Usability Context Analysis

6.1. Introduction

Usability is a classic concept of human computer interaction, described by Brooke (1998) as a general quality of the *appropriateness to a purpose* of an artefact. This means the usability of any tool or system has to be viewed in terms of the context in which it is used, and its appropriateness to that context.

This chapter discusses contextual analysis for the development of virtual reality applications, applied to safety training in mines. The purpose of this contextual analysis is to obtain information that will contribute to the usability of systems under design. It also contributes to the design of subsequent usability evaluations. The results of the contextual analysis were applied to the design of a prototype, which was used and evaluated at a large platinum mine. This evaluation was part of Cycle 1, as described in Section 5.6 of the chapter on research design.

This section refers to usability context analysis (UCA). It can also be termed contextual analysis, as in the previous paragraph. The term *context of use* is a related concept, which refers to the objects, tasks and environment of a system.

Section 6.2 provides a background to usability context analysis, with reference to the ISO 9241 standard. Section 6.3 provides examples of UCAs from the literature, while Sections 6.4 and 6.5 deal respectively with the application and findings of the UCA conducted at the platinum mine. The next sections, Sections 6.6 and 6.7, give details on the prototype, *Look, Stop and Fix*, and reports on the findings of an early evaluation.

It is important to note that the research methods and findings discussed in this chapter are not the main research of this study, but are only relevant to the UCA as part of the first cycle of the research design. The findings of the UCA led to the development of the first prototype, also discussed in this chapter. The main research findings of this study are presented in Chapters Eight and Nine.

Much of the material of this chapter was included in two conference papers. The first, by Van Wyk and De Villiers (2008), was presented by the researcher at the annual conference of the South African Institute of Computer Scientists and Information Technologists in 2008 and was published in the conference proceedings. The conference paper is attached as Appendix D-1, and its content influenced Sections 6.2 to 6.5. The second paper, Van Wyk and De Villiers (2009), was presented by the researcher at the ACM Afrigraph Conference in 2009 and was published in the conference proceedings. This conference paper is attached as Appendix D-2, and its content influenced Sections 6.5 to 6.7. The work in both the papers was done specifically for the purpose of this doctoral study.

The layout of this chapter is shown in Figure 6.1. Figure 6.2 indicates which part of the research process is covered by this chapter. The coloured sections of DBR Cycles 1 and 2 are relevant to this chapter, that is, the complete Cycle 1 and the first three phases of Cycle 2. In Chapters Seven, Eight and Nine there are similar figures to Figure 6.2, indicating the topics addressed in each chapter.

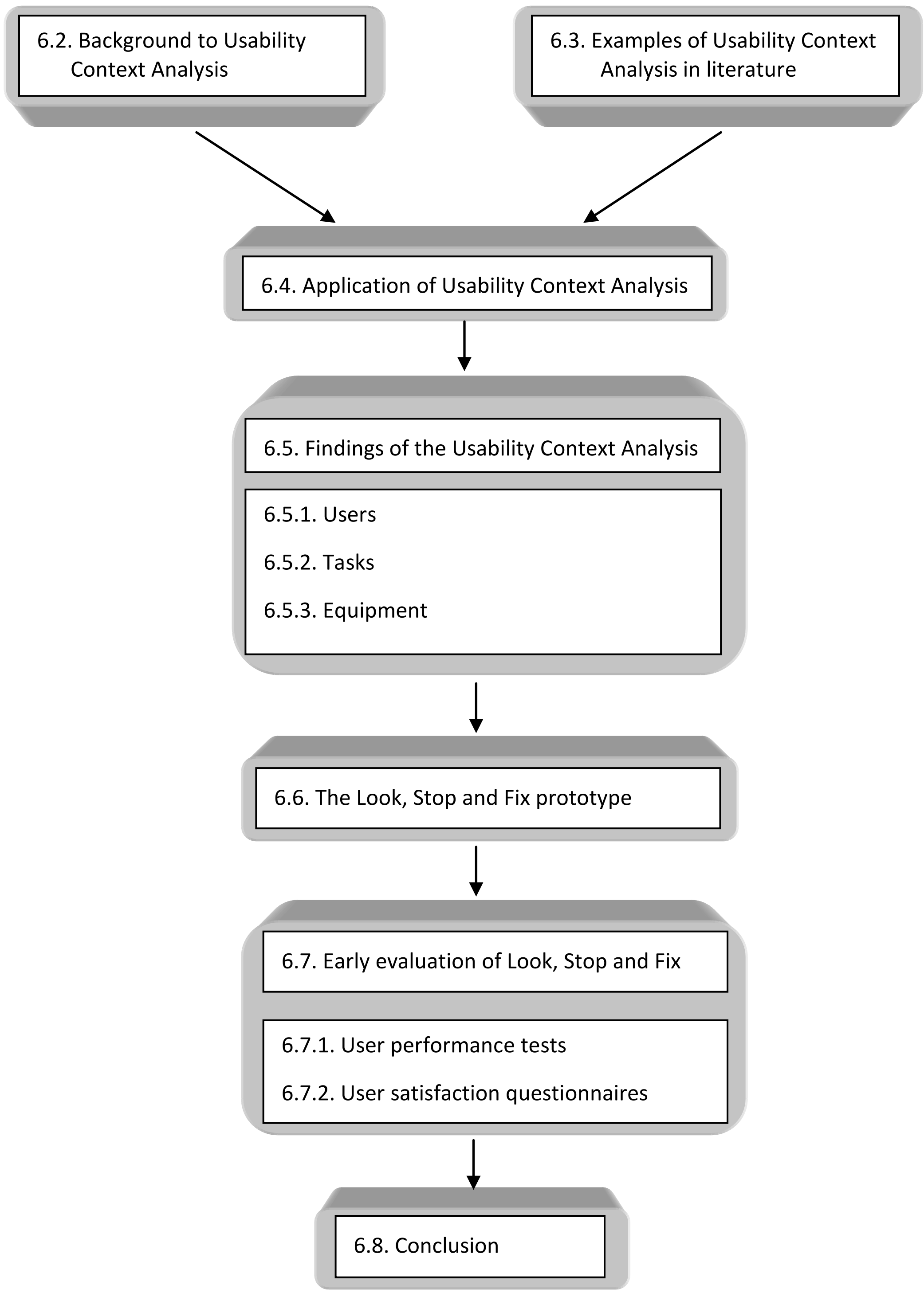


Figure 6.1: Layout of Chapter Six.

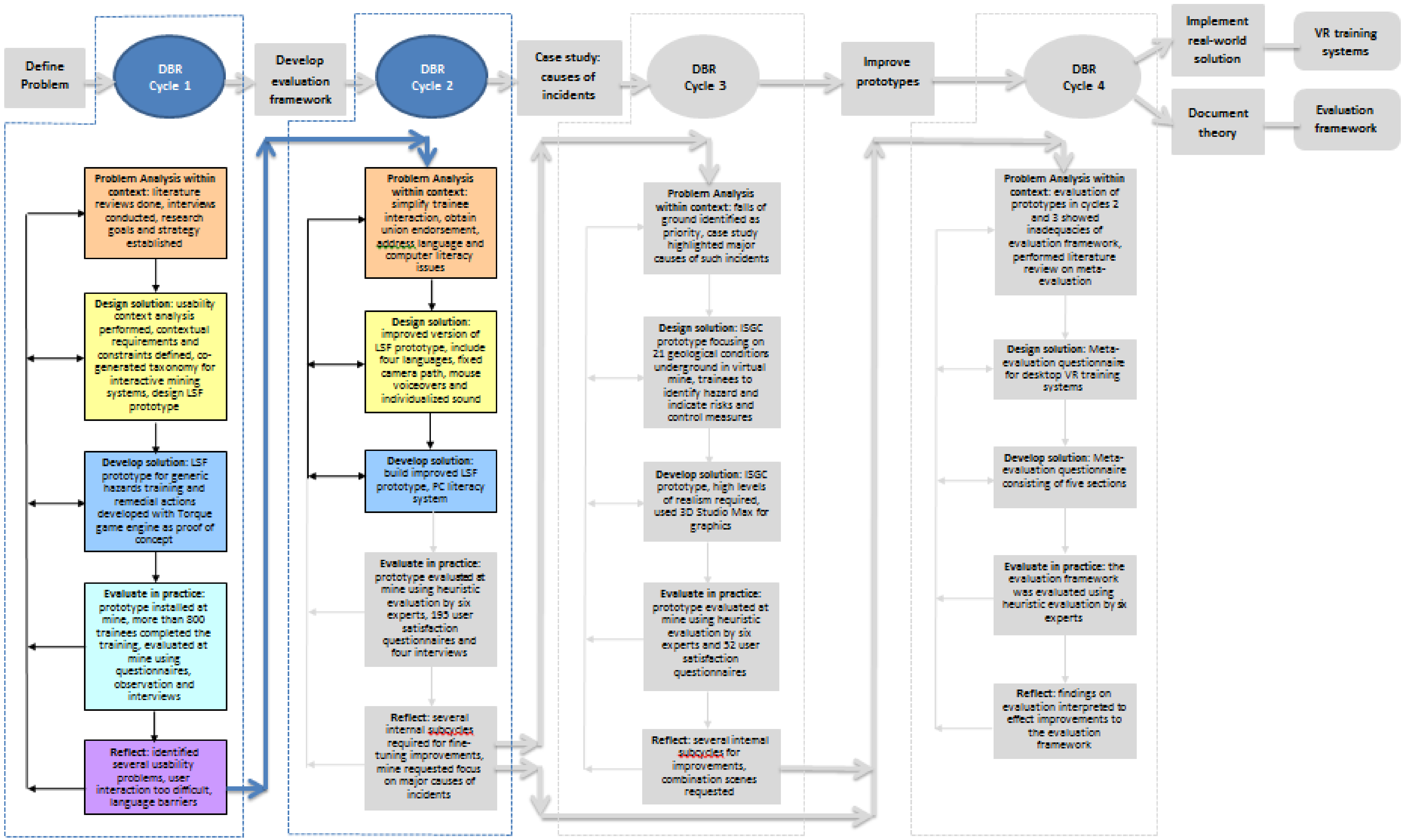


Figure 6.2: Research process flow discussed in Chapter Six.

6.2. Background to Usability Context Analysis

The International Organisation for Standardisation defines usability as the "extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO 9241-210, 2010). According to this ISO, the context of use includes the users, tasks, equipment (hardware, software and materials), and the physical and social environments in which a product is used.

6.2.1. Usability Context Analysis and context of use

When a system has been developed, it will be used within a particular context. According to Brooke's (1996) classic work, it is required to first define the intended users of the system and the tasks that they will perform with it. Furthermore, it is also necessary to define the characteristics of the physical, organisational and social environments in which a system will be used. Without specifying these contextual aspects, it is not possible to specify the usability of a system.

Context of use is incorporated into the ISO 9241 standard on the ergonomics of human-system interaction (ISO 9241-210, 2010). This ISO defines the process of understanding and specifying the context of use as one of the main stages within the human-centred design process. Figure 6.3 depicts the stages of the human-centred design process as shown in ISO 9241-210. The figure is followed by an explanation of each stage.

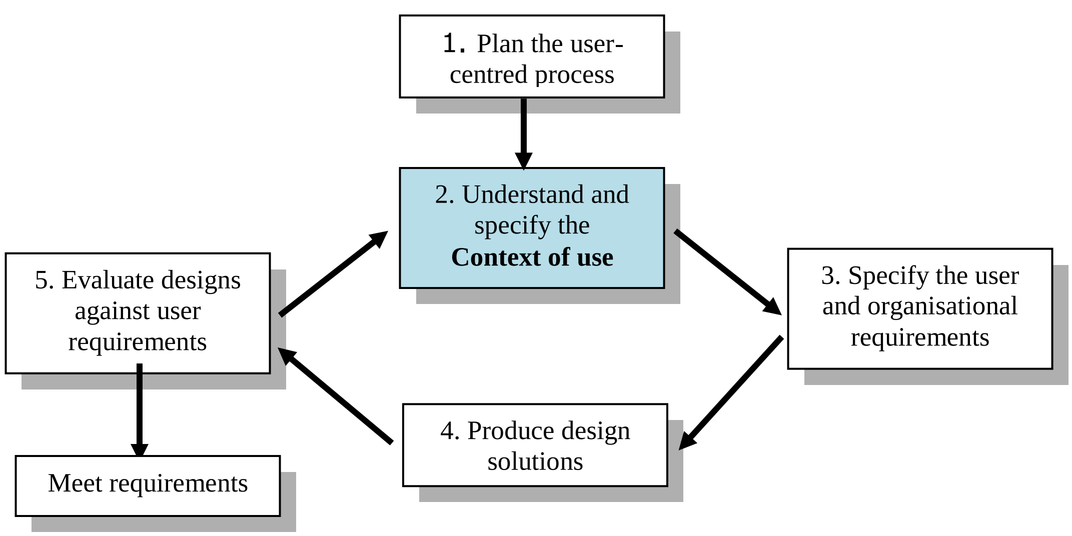


Figure 6.3: The human-centred design cycle (from ISO 9241-210).

6.2.2. Usability Context Analysis and human-centred design

The ISO 9241-210 standard provides a framework for human-centred design activities, comprising five stages shown in Figure 6.3. The second, third, fourth and fifth stages are explicitly joined in a loop to ensure iterations until the design objectives have been achieved. The goal of the design cycle in the framework is to design a system that effectively meets the user requirements. The stages are as follows:

- **Plan the user-centred process**: This first stage requires getting buy-in to the user-centred design philosophy from all the stakeholders involved in the development process. Most importantly, it involves the development of a plan for eliciting the user requirements and for testing.
- **Understand and specify the context of use**: The quality of use of a system depends on the extent of understanding and planning for the characteristics of the users, the tasks and the organisational and physical environment in which the system will be used. Rogers, Sharp and Preece (2011) refer to context of use as relating to four aspects of environmental requirements, namely: the physical, social, organisational and technical environments. It is important to understand and identify the details of this context in order to guide early design decisions, and to provide a basis for specifying the context in which usability should be evaluated.
- **Specify the user and organisational requirements**: This step involves specifying the functional requirements for the system, as well as the organisational requirements and the needs of the user in relation to the context-of-use description.
- **Produce design solutions**: Potential design solutions are proposed and prototypes are developed. Users are exposed to the prototypes and are studied as they perform specified tasks. Feedback from use of the prototypes is used to improve the design and this process is iterated until the design objectives are met.
- **Evaluate designs against user requirements**: Formative and summative evaluation methods are used to improve designs and to assess whether user and organisational objectives have been achieved.

Usability Context Analysis (UCA) is thus a structured method for eliciting detailed information about a product and how it will be used, and for deriving a plan for a user-based evaluation of the product. In this method stakeholders should meet to detail the actual circumstances (or intended use) of a product (Barisic, Amaral, Goulao, & Barroca, 2012; Hay, Kim & Roy, 2005; Mills, 2007; Van der Linde, Wessels & Kirakowski, 2013).

6.2.3. Usability Context Analysis and usability evaluation

With the ISO definition of usability given at the start of Section 6.2 as a background, the ISO 9241-11 standard provides a framework for specifying and evaluating usability in terms of user performance and satisfaction. This framework is depicted in Figure 6.4 (ISO 9241-11, 2008). User performance is measured by the extent to which the intended goals of use are achieved (effectiveness) and the resources that have to be expended to achieve the intended goals (efficiency). Satisfaction is measured by the extent to which the user finds the use of the product acceptable.

ISO 9241-11 emphasises that usability is dependent on the context of use and that the level of usability achieved will depend on the specific circumstances in which a product is used. The context of use consists of the users, tasks, equipment, and the physical and social environments, as indicated in the left block of Figure 6.4. All of these may influence the usability of a product.

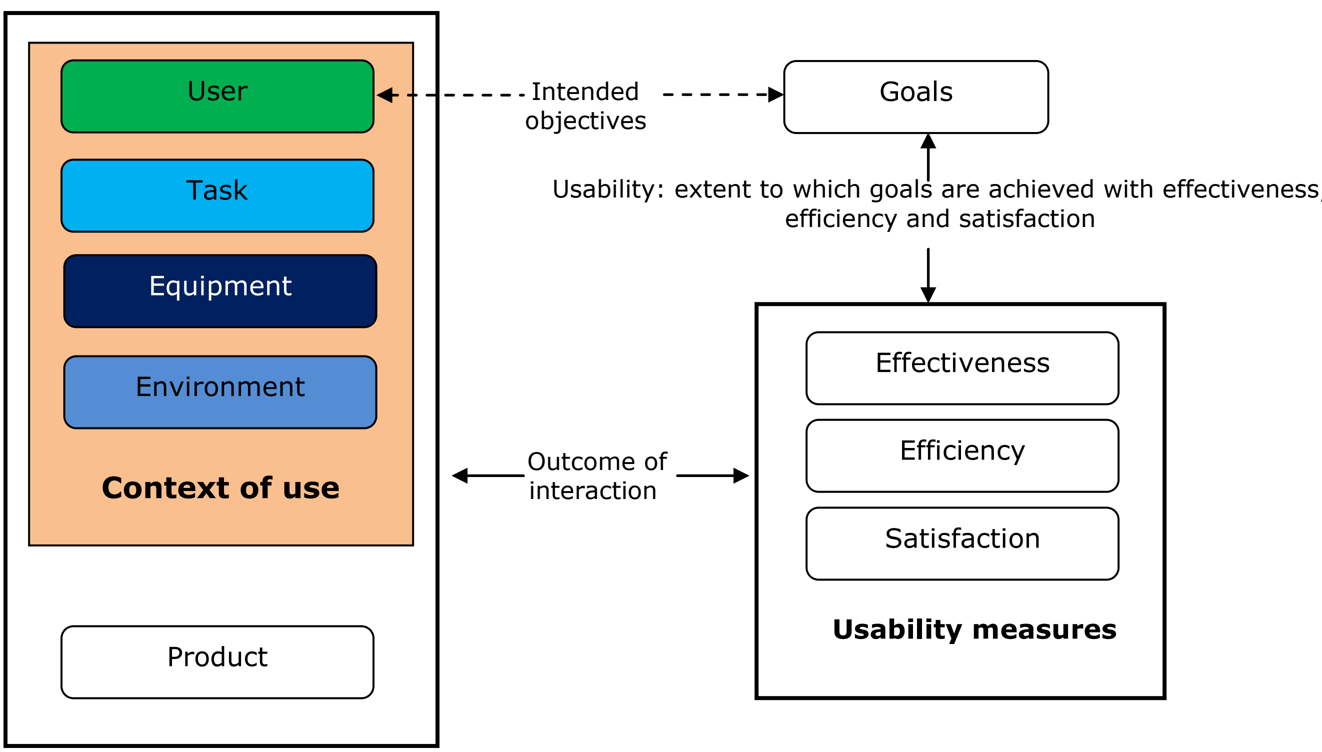


Figure 6.4: Usability Framework (from ISO 9241-11).

In the present study, the methodology of UCA has been extended to include all the context-of-use aspects of the ISO 9241-11 usability framework. Usability context analysis provides a framework to ensure that all factors which may affect the usability of a product are considered. The purpose of the context analysis described in this chapter, in Sections 6.4 and 6.5, is to contribute to the usability of the system designs presented in this research. It also provides a basis for subsequent design of usability evaluations. According to INUSE (2004), conducting a UCA helps to reduce the number of assumptions made by the analyst. It also ensures that the client is involved in the design process and that factors affecting usability are considered before the design stage.

According to Cramer, Evers, Zudilova and Sloot (2004:177), "Virtual reality (VR) applications are often developed relatively independently from the real contexts in which they are going to be used". Performing a UCA for VR training systems in the mining industry is therefore pertinent in the present study, which aims to guide the development of future VR training systems, as well as to answer questions about the real-life tasks that are undertaken, the context that these systems should support, functionality of such systems and their level of immersion. Research Subquestion 3, defined in Chapter One, is addressed through the UCA, namely: "What are the contextual requirements for virtual reality training systems for the mining industry?"

6.3. Examples of Usability Context Analysis in the literature

Maguire (2001:453) states that "when assessing a product from a Human Factors point of view, there is a tendency to forget about the Context of Use". Information Technology products are often divided into those which are usable and have satisfactory ergonomic features and those which are not. It is argued that it is incorrect to describe a product as ergonomic or usable, without also describing the *context* in which it will be used. Maguire points out that the UCA should consider for whom the product was designed, for what it will be used, and the environment in which it will be used.

Maguire also proposes a process for performing a context analysis. The method is particularly aimed at non-experts in the domains of user-centred design and evaluation. This process, which stresses the value of conducting context-of-use analysis at all stages of design and development, was applied to an automatic teller machine (ATM) example. In this example it was highlighted that reconciling user requirements with technical and business requirements is a complex issue. The advantage of applying context-of-use analysis throughout the design lifecycle is that it forms a complementary strategy that addresses both user requirements specification (early on) and user-based testing (at later stages). It is concluded that "an understanding of the context of use forms a useful input to the process of specifying usability requirements, constructing a design prototype which can be evaluated, and evaluating the prototype with typical end-users" (Maguire, 2001:481).

Mills (2001) reports on the important role of task analysis as part of a UCA to ensure fitness for purpose. By using an example of an echosounder, Mills indicates that a task-based approach can highlight discrepancies within usage. Another article by the same author advocates conducting UCAs, stating that "studies involving a critical assessment of usability context analysis within the software domain are scarce" (Mills, 2007: 499).

Mosqueira-Rey, Alonso-Rios and Moret-Bonillo (2009) highlight the fact that, even though context of use is widely recognised to be important, they have found no model that clearly describes all its features in detail. They propose a detailed taxonomy that defines context-of-use attributes relevant to usability, with the main attributes in the taxonomy being the users, tasks and environment. This is in line with ISO 9241-210 (2010), as discussed in Section 6.2. Other examples of context-of-use studies in literature include analysis of the context of use for a training application for oil refineries (Träskbäck & Haller, 2004), context-of-use components for mobile systems (Coursaris & Kim, 2007), and a context-of-use classification for automated translation systems, defined in terms of characteristics of the translation task, input characteristics, and user characteristics (Estrella, Popescu-Belis & Underwood, 2005). Hörold, Mayas and Krömker (2012) report on a usability context analysis done on the German public transport system. Nation-wide case studies were analysed in terms of users, tasks and the environmental context in order to conclude usability requirements for passenger information systems.

Alonso-Rios, Vazquez-Garcia, Mosquiera-Rey and Moret-Bonillo (2010) state that the results of usability studies of products in specific usage contexts cannot be directly generalisable to other environments. Stressing a similar point on varying contexts, however, Brown, Sharples and Harding (2013) describe a usability evaluation process for geographic information, which is based on a structured framework that allows for varied context of use.

Limited research has been published on usability context analysis in VR. Cramer, Evers, Zudilova and Sloot (2004) discuss contextual analysis in VR applied to the case of a virtual radiology explorer (VRE) system aiding medical diagnosis and planning. This VRE prototype was developed in response to needs articulated within the medical world for research into VR visualisation and simulation of physiological properties. Cramer *et al.* suggest that solutions for potential usability problems cannot be found without involving prospective end-users and obtaining more detailed knowledge about the VRE's context of use. They note the benefits of contextual analysis, particularly the role it played in generating information that supported and enhanced the development of the VRE (Cramer *et al.*, 2004:185).

6.4. Application of Usability Context Analysis

In applying UCA in this study, the approach was to conduct some of the context analysis prior to designing the technological interventions, and also to obtain further data later by studying a prototypical VR system in its environment of use.

In DBR Cycle 1 of this research (Section 5.6.7 in Chapter Five), data collection was done in an integrated fashion, addressing several of the context-of-use issues in each of the data collection instruments or sessions. The following methods were used:

- Semi-structured interviews with three mine managers, four safety, health and environment officers, and three mine training managers.
- Structured interviews with 23 randomly selected mine workers.
- Informal observations of current training methods at five different mines.
- Questionnaires completed by trainees including questions on specific context-of-use issues.
- Informal observation at several underground mining stope areas to observe miners performing their daily tasks.

The interviews and observations were used at an early stage to investigate the environment and the target group, while the questionnaires addressed general contextual issues that emerged from use of the prototype.

The researcher carried out informal observations in an unobtrusive manner, observing participants while they were working underground. Questions were asked to verify the information gathered during the interviews before or after work activities. Photographs and video material were taken whenever possible and where permissible.

These studies are described in Section 6.5. The findings of the studies contributed to the design of a prototype, called *Look, Stop and Fix*.

The subsequent evaluation of the prototype is discussed in detail in Section 6.7. It was evaluated using the following methods:

- Questionnaires completed by 221 trainees after completion of the prototype training system.
- Structured interviews conducted with 23 randomly selected mine workers.

6.5. Findings of the Usability Context Analysis

This section discusses in part Research Subquestion 3, while the case study to be presented in Chapter Seven further supplements the contextual requirements mentioned in this section.

RQ3	What are the contextual requirements for virtual reality training systems for the mining industry?

In addressing this subquestion, the UCA aimed to investigate the contextual requirements and constraints for VR training systems for the mining industry. As indicated in Figure 6.4, the major context-of-use issues are users, tasks, equipment and environment. The findings that follow are presented under these headings. In the case of environment, a distinction is made between the workplace environment and the training environment.

6.5.1. Users

The South African Department of Minerals and Energy uses the Mining Industry Standard Code of Occupations (Department of Minerals and Energy, 2008), listing 1032 job titles in the mining profession. For the purpose of this study, only underground mine workers were observed and interviewed, involving mainly the following jobs: belt attendant, miner, cheesa (miner's assistant), rock drill operator, loco driver, panel operator, shift supervisor, team leader, stope timberman and winch operator.

A total of 23 structured interviews were conducted with randomly selected workers, as part of the UCA study. The interviews were conducted by clerks at the mine in each interviewee's preferred language and their answers were transcribed on an interview template. This section presents data relating to the profile of the user group, while workplace-related findings are given in Section 6.7.3. Analysis of the interview responses yielded the following results related to human-resource issues:

- The subjects had different cultural backgrounds and spoke various languages.

- Some had a very limited understanding of English.
- As indicated in Table 6.1, ages were between 20 and 60, with an average age of 36.4.
- Levels of education varied from Grade 5 to Grade 12. The majority had secondary education, twelve of them (52%) having matriculated or at least commenced matriculation level studies. Only two (9%) had not done any secondary studies (see Table 6.2).
- The interviewees had various levels of underground mining experience, ranging between 2 years and 25 years (see Table 6.3).
- All the interviewees were men. At the time of the interviews, most underground workers were men, but some mines are appointing women in certain of these positions.
- The majority of the interviewees (74%) were confident that they could perform their duties well.
- More than 80% of the interviewees had never used a computer.
- To determine usage of technology, workers were asked whether they used cell phones and ATMs. For both these technologies, more than 80% answered in the affirmative. Despite the wide range in literacy level and minimal computing experience, they were not afraid of the prospect of computer-based training. In fact, they were of the opinion that they would enjoy it!

Table 6.1: Ages of the interviewees.

Age Group	21 – 30	31 – 40	41 – 50	51 – 60	Total
Number of Participants	8	7	6	2	23
Average age = 36.4					

Table 6.2: Highest grade at school completed by the interviewees.

School Grade	Grade 5 – 6	Grade 7 – 8	Grade 9 – 10	Grade 11 – 12	Total
Number of Participants	2	3	6	12	23

Table 6.3: Interviewees' experience in the mining industry.

Years of Mining Experience	0 – 5	6 – 10	10 – 15	16 – 20	21 – 25	>25	Total
Number of Participants	6	4	7	4	2	0	23

It also emerged from the interviews that workers are concerned about safety and the high number of accidents in the industry. The researcher noted that the National Union of Mineworkers had organised a formal protest march in December 2007 in which thousands of mineworkers had marched in central Johannesburg, urging management to improve safety conditions. Concern is also raised by the production bonuses offered by certain mines, which might place the emphasis on production at the potential cost of safety.

6.5.2. Tasks

Many varying tasks undertaken by different categories of underground workers were identified in observation and in discussions with the Safety, Health and Environment (SHE) managers and workers. Miners are exposed to a number of risks, but for the purpose of this study, it was decided to focus on *hazard recognition*, *identification* and *correct procedures in addressing hazards,* as explained in Section 1.9.1. All underground workers should be aware of hazards in the workplace, both in the haulage and in the stope area.

Creation of awareness of hazards and their consequences in the workplace holds benefits for health and safety, as it helps to prevent accidents that cause loss of human lives, production loss and lower morale in the industry.

Hazards can be classified as generic or job-specific. All underground workers should be able to identify and fix generic hazards, while job-specific hazards refer to potentially dangerous conditions that can occur while miners perform duties related to a specific job or role. For example, while working with the winch, a winch operator may encounter various hazardous conditions to which a rock drill operator may not be exposed. These are job-specific hazards. However, since the winch cables and snatch blocks are used within the stope area, there are certain generic winch hazards of which all workers should be aware.

Analysis of the data obtained via interviews and observation led to the categorisation of generic hazards into five groups:

1. Employee actions
2. Geological conditions and support
3. Machinery and equipment
4. Poor house-keeping
5. Sub-standard conditions.

For each of these task groups, workers should be able to recognise particular hazards, identify them correctly and follow the correct procedure in dealing with them.

6.5.3. Equipment

In the observations underground, the researcher noted that the equipment used by workers depends on their particular duties. Photographs taken by the researcher are used in this section to illustrate appropriate equipment. For example, a cheesa might use a pinch bar to make an underground workplace safe by barring down loose rocks (see Figure 6.5), whereas the rock drill operator uses a pneumatic or hydraulic rock drill. It was mentioned by SHE managers that specialised training is required for using certain tools and equipment.

Workers are required to wear the correct PPE (Personal Protective Equipment), which usually includes a hard hat, appropriate belt and overall with reflective strips, boots, ear plugs, gloves, cap lamp, battery pack and protective glasses. Incorrect behaviour in terms of not wearing the correct PPE is also a generic hazard. Although many areas underground are well illuminated, a cap lamp serves just as much for others to notice a worker as it does for that worker to see properly. The cap lamp should be switched on at all times when workers are underground.

Figure 6.5: Worker using pinch bar to dislodge loose rocks.

The following types of machinery are frequently found in underground mines:

- Winches,
- Locomotives, hoppers and flat cars,
- Axial ventilation fans,
- Pumps,
- Drilling equipment,
- Continuous miners,
- Scrapers, tips and chutes for handling rock,
- Chairlifts (see Figure 6.6),
- Conveyor belts, and
- Trackless mining machinery, e.g. load haul dumpers, drill rigs and utility vehicles.

In general, machines can become very hot when used, might produce very high pressures, and some can be moved around. Because of their size and mass, any inadvertent contact with machinery poses a real and immediate danger to workers.

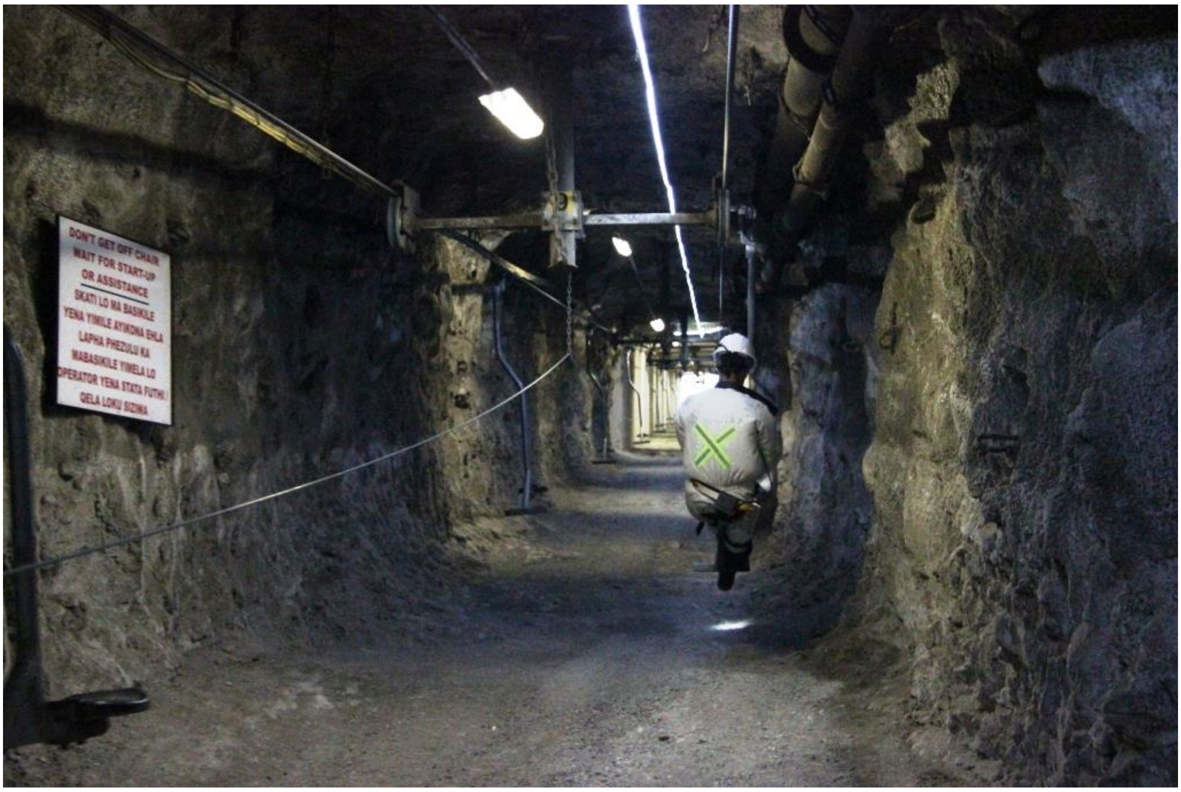


Figure 6.6: A worker riding on a chairlift to the next mine level.

6.5.4. Work environment

Based on observations, the underground work environment can be described as dirty, dark, wet, noisy, hot, uncomfortable and dangerous. Hazards related to the work environment can include the following:

- Working in confined areas.

- Working in steeply inclined excavations.
- Handling heavy material and equipment.
- Working in the proximity of moving machinery.

The sizes of the stope areas, where drilling and blasting occur, vary from one mine to another, but the stopes observed in the Mine in this study were approximately 27 metres wide with a maximum height of 1.5 metres. The gulley area next to the stope, from which the scraper winch extracts the blasted rock, is about 2.5 metres high, allowing workers to stand upright when not working in the stope. Rock drill operators work in a sitting position while drilling holes for the explosives. The use of narrow stopes is due not only to the reef usually being 1metre wide or narrower, but also because smaller excavations are inherently safer than bigger ones. This leads to the same excavations being used to transport material into the stope and rock out of it, as well as to provide access for workers to and from the working face. Workers need to bear in mind that their working space is also used for a variety of other purposes, many of which can cause injury. Figure 6.7 shows workers taking a rest in a typical underground working area.

Figure 6.7: A typical underground work environment.

Blasted areas are watered down as part of the cleaning process, causing the areas to be wet. Excessive water is hazardous and can lead to slip-and-fall incidents. Conditions underfoot vary and are often very slippery. It is extremely important for workers to watch carefully where they walk, because they may step into drain holes. When walking through water and unable to see the footwall, one needs to walk very slowly, feeling for each footing before transferring one's weight. Furthermore, the workshops often have heavy machinery and oily, slippery floors. Care should be taken to be aware of falling objects and precautions should be taken not to slip and fall in a workshop area.

Access to the place of work is often via a main transport haulage, also used by locomotives and other vehicles. Pedestrians should always stand still in a safe area when a vehicle is passing and should walk only in designated travelling ways. There is usually a waiting place close to the entrance to the working place. Areas after the waiting places are often subject to regular removal of rock, either by blasting or by continuous mining. This necessitates that these areas should be inspected and made safe at the beginning of each shift and at regular intervals during the shift.

In a typical hard rock mine, direct access to a stope area occurs via a cross-cut. This is also a congested area with material being transported and rock being removed. Dangerous areas are barricaded off with hazard tape. Winches in centre gullies move rock to ore passes. The top of an ore pass is covered by a grizzly so that workers cannot fall into the ore pass. Handrails, chains or other barricades are often erected around ore passes, to prevent workers falling into them.

Safe storage, handling and use of explosives is essential. Moreover, when explosives detonate, nitrogen and carbon dioxide are released. While these are not poisonous, they can result in asphyxiation since they dilute the oxygen in the air. In addition, nitrous fumes and carbon monoxide, which are poisonous, are produced in high concentrations immediately after a blast. Furthermore, the detonation of explosives creates large quantities of dust which should be allowed to settle before workers are exposed to the area.

With ground falls being the main cause of fatalities in the industry, it is essential to correctly identify different geological conditions, especially after blasting. Conditions such as shear zones, joints and dykes should be supported correctly to prevent falls of ground, and loose rocks should be barred off. A high proportion of rockfall accidents occur during re-entry after blasting, when the initial inspection and making-safe procedures are conducted to stabilise the rock before work recommences. Geological conditions are discussed in more detail in Chapter Seven.

Despite proper management and maintenance of underground areas, conditions are dynamic and can change without warning. Each miner must have insight and be acutely aware of hazards and potential hazards.

Generic workplace hazards usually relate to support conditions, ground conditions, inadequate escape ways or obstructions in escape ways, fire, exposure to unsafe electrical connections, humans in proximity of an area where loose rocks are to be barred, and working under unsafe roofs or sidewalls.

These findings were corroborated by data collected from interviews with safety and health managers.

6.5.5. Training environment and organisational aspects

Training is done in accordance with the unit standards specified by the Mining Qualifications Authority. Most mines have training centres where new recruits are trained in job-related courses. This training is predominantly instructor-led and occurs in a class-based environment. In general, the use of technology in training is limited, but some mines do have computer-based training facilities. Classroom training is usually followed by practical training in the real work environment, until the instructor certifies the trainee as competent to perform the work correctly and safely. For this purpose many mines have underground training areas (see Figure 6.8).

Figure 6.8: An underground training area at a platinum mine.

Workers returning from annual leave, so-called 'ex-leaves', go to the training centre for refresher courses before being permitted to return to the underground milieu. A factor of concern noted by the researcher is the lack of assessment after these refresher courses for ex-leaves. Sometimes workers merely sign a form indicating that they have worked through the files.

The issues identified in this integrated data collection process, as described in Sections 5.6.6 and 5.6.7, posed a challenge to the researcher in his role as potential designer of a VR prototype for focused training to address the problems related to hazards. It emerged from interviews with mine managers and mine health, safety and environment officers that improved and technologically advanced training systems are required to assist mine management and SHE management to improve the safety records at their mines (Baker, 2006; Lubbe, 2006; Moldenhauer, 2006; Stander, 2006; Wenhold, 2006).

Section 2.2.2 discussed three categories of VR systems, namely fully-immersive, semi-immersive, and desktop or non-immersive systems. To accommodate high volumes of trainees, the use of desktop VR systems on ordinary personal computers can provide an affordable means of achieving current training goals in the South African mining industry. With this in mind it was decided to develop a desktop VR prototype focusing on generic hazards. The system would be delivered to ex-leaves as refresher training and would assess their abilities to identify and rectify hazards.

6.6. The Look, Stop and Fix Prototype

In support of the findings of the context-of-use analysis for mine safety training, a virtual reality interactive training system was designed and developed. This prototype, called *Look, Stop and Fix,* focused on hazard recognition and training in remedial actions in conventional mining.

The prototype simulates the underground working areas, incorporating potential hazards that mine workers need to identify. Trainees have to spot potentially hazardous conditions, identify the hazards correctly, and indicate which actions should be taken to address the situation. Failure to correctly identify a hazard or to specify the correct action in dealing with such a hazard causes an animation to play out, displaying the possible disastrous consequences of ignoring or incorrectly responding to such a hazard.

The context-of-use analysis provided valuable information to inform the design of the prototype. Based on this information, it was decided that trainees would be placed in a

3D virtual underground environment, where the haulage and stope areas would be realistically simulated. Various generic hazards would be present in this environment, selected randomly from the five categories of hazards generated from the UCA, as discussed in Section 6.5.2. Table 6.4 lists the 27 hazards included in *Look, Stop and Fix*, and also indicates which hazard category each hazard belongs to, according to the following category classification:

1. Employee actions
2. Geological conditions and support
3. Machinery and equipment
4. Poor house-keeping
5. Sub-standard conditions.

Table 6.4: Generic hazards included in the prototype.

Hazard	***Category***
H01 : Rail gates not in position	***5***
H02 : Material stacked incorrectly on the material car	***4***
H03 : Loco blocks entry to working area	***3***
H04 : Short on electrical cable at substation	***5***
H05 : Open explosive box	***4***
H06 : Loose rock falls on miner's head	***2***
H07 : Employee working in haulage gets run over by the loco	***3***
H08 : Employee falls over materials stacked incorrectly	***1***
H09 : Dangerous sidewall conditions	***2***
H10 : Mud-rush in haulage	***5***
H11 : Damaged support in waiting area	***1***
H12 : Pipes/Material stacked on the loco railway	***4***
H13 : Material is lying in the passage	***4***
H14 : Some segments on grizzly are missing	***5***
H15 : No drum guards on winch	***3***
H16 : Bad hanging wall	***2***
H17 : No barricade or chain around grizzly	***5***
H18 : Rigging not safe, snatch block hits miner	***3***
H19 : Cover on lockout device is broken off	***4***
H20 : Not enough support at interconnecting brow and slip	***2***
H21 : Footwall area not cleared, misfire of explosives	***1***
H22 : Water spray hazard	***1***
H23 : Water in the area where people should walk	***4***
H24 : Excessive water	***4***
H25 : Employees are horse playing in the mine	***1***
H26 : Material stacked incorrectly in the haulage	***4***
H27 : Inadequate support in working area	***2***

In the *Look, Stop and Fix* prototype, a camera pans slowly through a virtual mine, following a predetermined path. At the bottom of the screen, a STOP button and a REPLAY button are visible. At any point in time, while the camera is moving through the mine, the trainee can use the mouse to click on any of the buttons. If the trainee clicks on REPLAY, the previous scene is replayed and the trainee has another opportunity to view the environment. If the trainee clicks on STOP, the camera stops and two other buttons appear, GO and IDENTIFY, as indicated in Figure 6.9. (Note: Figures 6.9 to 6.18 are screenshots extracted from the prototype. The screenshots may appear dark. This is an accurate reflection of the actual underground work environment and the VR prototype was developed to portray realism).

Figure 6.9: Screenshot of available options when trainee has stopped the simulation.

If, after stopping, the trainee feels that there is no hazard, then clicking on GO will continue the journey forward. However, if the trainee is of the opinion that a hazard is indeed present, then the IDENTIFY button should be clicked. Names of various possible hazards are then shown as overlays and the trainee must select the appropriate one (see Figure 6.10). In the case of the correct hazard being selected, then a number of remedial actions are shown and the trainee should select the correct action. Figure 6.11 is an example of a situation where the trainee has to select the correct procedure. The simulation is in the form of a game and trainees receive a score for each correctly identified hazard, as well as for correctly indicating the procedure to deal with each hazard.

Figure 6.10: Example screen shot of options the trainee can select when identifying a hazard.

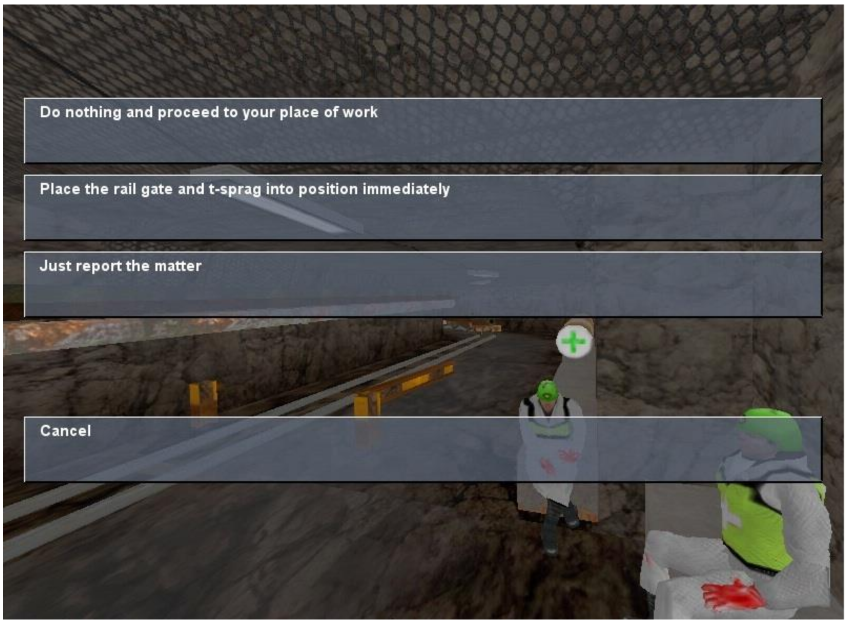


Figure 6.11: Example screen shot of options the trainee can select when specifying the correct procedure to deal with the identified hazard.

Different types of hazards have different score values. If trainees do not correctly identify a hazard or the correct procedure for dealing with such a hazard, an animation plays out displaying the possible disastrous consequences of ignoring it. See Figure 6.12, which portrays a worker accidentally pushing a material car down the shaft, due to a rail gate not being closed.

Figure 6.12: Series of screenshots depicting animation of the potential consequence of not closing the rail gate.

In the *Look, Stop and Fix* prototype, trainees select their language of choice for use in the system. The options are English, Tswana, Sepedi and Xhosa, as shown in the login screen in Figure 6.13.

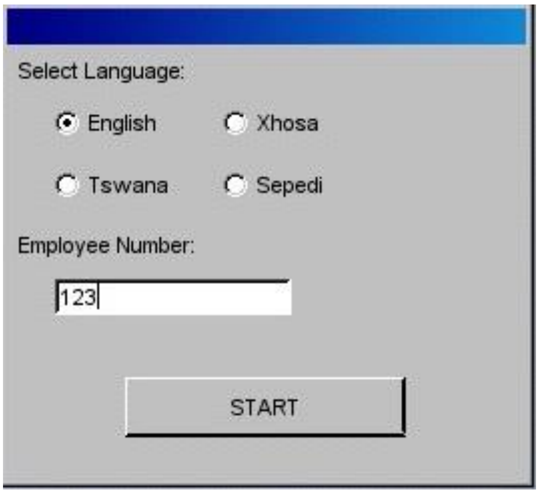


Figure 6.13: Language options available when logging in.

The selected language is used throughout the simulation for any textual feedback, as shown in Figure 6.14, which displays the Tswana version of the text interaction used for identifying the same hazard depicted in Figure 6.10, where the text is in English. Using audio mode via earphones, trainees can also listen to the text in the language of their choice when the mouse cursor moves over the written text. Furthermore, all the feedback and explanations are provided in the selected language.

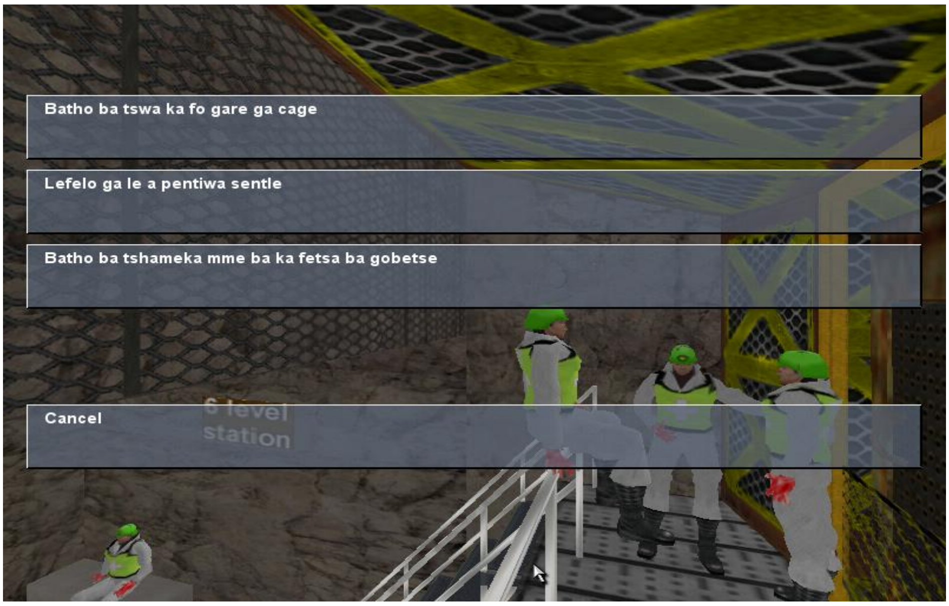


Figure 6.14: Example screen shot of Tswana interaction.

Figures 6.15 to 6.18 are further examples of screen shots from *Look, Stop and Fix.* They depict hazards at the workplace and the possible consequences of not taking remedial action to address these hazards.

In Figure 6.15 a trainee views an open explosive box. As this is a dangerous situation, the trainee needs to identify the hazard and specify the correct procedure for dealing with this situation. Figure 6.16 shows an animation of the possible consequence of ignoring this hazard.

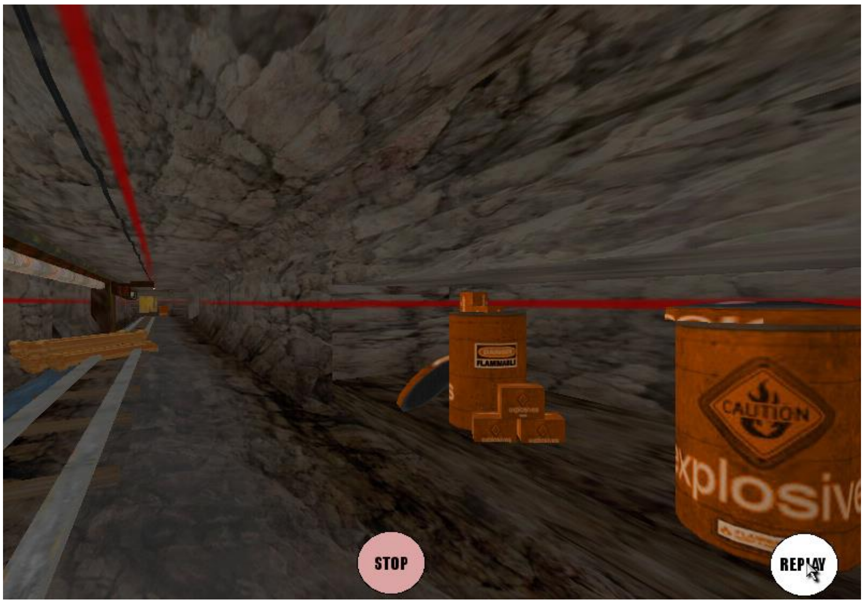


Figure 6.15: Screen shot of the Open Explosive box hazard.

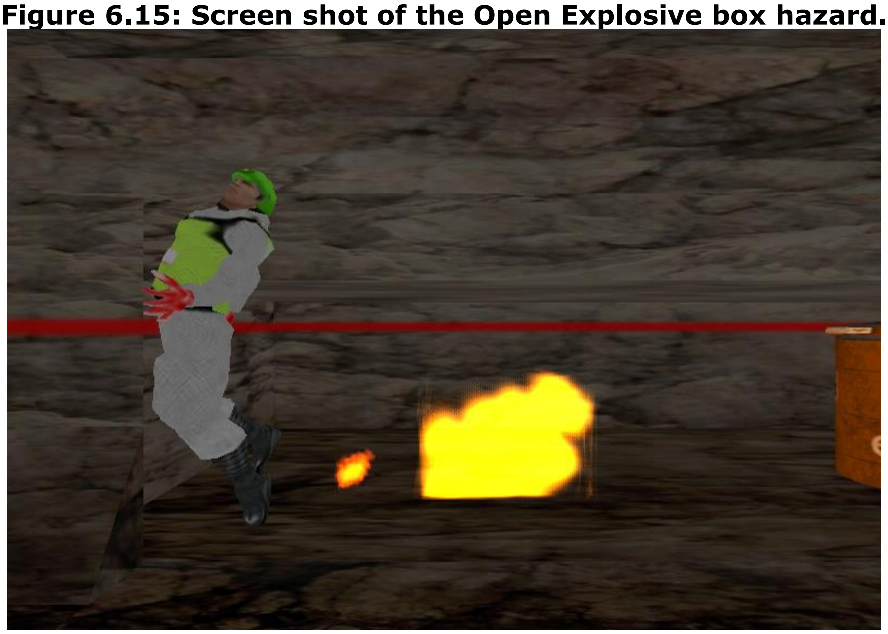


Figure 6.16: Animation of potential consequence of ignoring the open explosive box.

Figure 6.17: No drum guards on the winch.

Figure 6.17 portrays a hazard in the stope area, where the drum guards are not attached to the front of a winch. These drum guards protect winch operators from being accidentally pulled into the winch. The possible consequence of ignoring this hazard is illustrated in Figure 6.18.

Figure 6.18: Miner being pulled into the winch as a possible consequence of operating the winch without drum guards.

After reaching the end of the 27 hazards the trainee receives a final score. As a default setting, the training facilitator can specify the pass mark required prior to the trainee logging in. If the required pass mark is not achieved on completion of the system, the trainee is automatically placed back at the beginning of the first hazard, and must redo the training. If the trainee still does not pass after two attempts, the training facilitator is informed accordingly and a decision must be made regarding additional training interventions.

6.7. Early evaluation of the Look, Stop and Fix prototype

As described in Section 3.3, usability evaluation of a product or system is an approach that focuses on how well users can learn and use a system to achieve their goals. It also refers to how satisfied users are with that process. It is important to note that it is the system being tested and not the user (Rogers *et al.*, 2011). The evaluation described on this section is an early investigation of the first version of *LSF.* A formal evaluation of the improved *LSF* prototype is presented in Chapter Eight.

This section first reports, in Section 6.7.1, on performance of trainees undergoing training in *LSF* and then discusses the usability evaluation of *Look, Stop and Fix* in Sections 6.7.2. Section 6.7.3 reports on interviews which followed up on the UCA by investigating how suitable *LSF* was for its environment and context. These evaluations involved a combination of user performance tests, user satisfaction questionnaires and structured interviews.

6.7.1. User performance tests

The tests, conducted with 221 trainees, as mentioned in Section 6.4, measured trainee performance on hazard awareness tasks in the prototype system.

Pre-training

As indicated in Section 5.6.7, the *LSF* prototype was installed in a training facility at No. 10 shaft at the Mine. From informal observation of trainees using the prototype it became evident that, for many trainees, their lack of experience in using computers caused them to struggle with interaction. A pre-training program was then developed, which utilises videos and practice exercises to ensure that trainees master the required computer literacy skills. The program is comprised of four exercises to enable trainees to develop these skills:

- Drag and Drop
- Click – selection of items and movement of cursor
- Movie clips – visual introduction to PC and keyboard functions
- Catch the rock – game to improve hand-eye coordination, as trainees click on falling rocks aiming to do so before they hit the ground.

The training facilitators and the researcher were pleasantly surprised by the speed with which the trainees mastered the pre-training, which proved to be highly effective for its purpose. Its success could possibly be attributed to the fact that many of them used cell phones and/or bank cards at ATMs.

Prototype training and performance

Participants in the study logged onto the *Look, Stop and Fix* VR prototype using their employee numbers. They first selected the language they preferred to use in the system.

While working through the prototype, participants received marks for three actions:

- correctly spotting a hazard,
- correctly identifying the hazard, and
- taking the correct action in dealing with such a hazard (fixing the hazard).

The letters S (spot), I (identify) and F (fix hazard) were used to indicate results. After completion of the training, the system showed results in terms of S, I and F values, a final percentage, performance on each hazard (numbered H01 to H27), as well as the number of guesses (clicks on the IDENTIFY button when there was no hazard present). The pass mark was 80%. Most of those who did not pass had another attempt at working through the system. The average score achieved by the 221 participants for all the first attempts was 73%, with 107 of them achieving the 80% pass mark. The average achieved for the second attempts by the 114 participants who had not attained the required 80%, was 79%.

Figure 6.19 indicates the first-attempt-performance of the participants with regard to the five hazard groups, categorised during the UCA and discussed in Section 6.5.2.

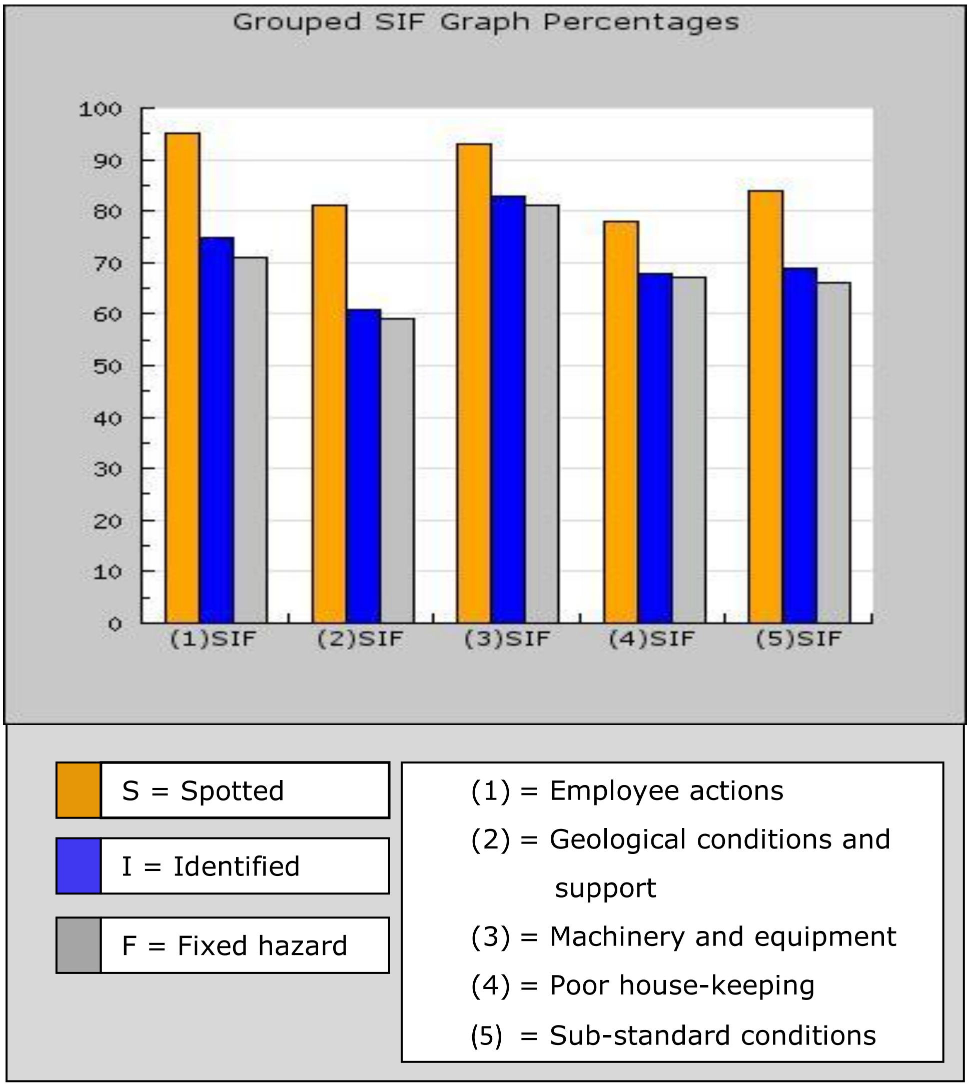


Figure 6.19: Spot, Identify and Fix percentages achieved per hazard group on first attempts.

Even though 95% of Group 1 hazards (unsafe acts by employees) were spotted correctly, the identification of hazards correctly spotted in Group 1 hazards and Group 2 hazards (geological conditions and support) proved to be more problematic than the other groups, with a 20% gap between hazards spotted and hazards correctly identified in both cases. Regarding identifying and fixing hazards, the participants performed best in identifying and fixing Group 3 hazards (machinery and equipment), 83% and 81% respectively. Trainees spotted only 78% of Group 4 hazards (poor house-keeping), the lowest of the four groups. For Group 5 hazards (sub-standard conditions), 84% were spotted correctly, but only 69% were correctly identified and 67% correctly fixed. In all five groups there was a very small difference between the number of hazards identified and the correct actions specified, indicating adequate knowledge of addressing hazards once they were correctly identified.

6.7.2. User satisfaction questionnaires

To evaluate user satisfaction of the prototype, a computer-based questionnaire was completed by each of the 221 participants after completion of the prototype training. Table 6.5 indicates the results of the questionnaire, presented as percentages.

Table 6.5: Results of computer-based questionnaire.

Aspect evaluated	Yes (%)	No (%)
Easily identify objects	87	13
Easy to use	93	7
Easy to understand	93	7
Enjoyed the system	95	5
Prefer VR training	84	16

From Table 6.5 it is clear that the computer-based questionnaire yielded the following major results, with associated usability aspects indicated in brackets:

- 87% of the participants indicated that they could easily identify all the objects in the simulated environment (high visibility). By 'objects' they were referring to modelled items in the virtual environment, such as winch pulleys, safety signs, support material.
- 93% found the system easy to use and understand (learnability and ease of use).
- 95% indicated that using the prototype was an enjoyable experience (user satisfaction).
- 84% indicated that they would prefer VR training to other types of training (method of choice).

The participants were also asked in open-ended questions whether they thought the hazards presented in the prototype system (authenticity) could really happen. Most participants felt that all the hazards portrayed were real hazards in their working environment. It is interesting to note that in a follow-up question as to whether the accidents in the system could actually happen to them, many thought that most of the accidents would not happen to them personally. They could have meant that certain hazards would not be a threat to them in their particular job and role, but it could also be an indication of laxity and an attitude that accidents were more likely to happen to others than to them personally.

6.7.3. Structured interviews

In order to obtain more information, structured interviews were conducted three months later with 23 randomly selected employees, who had all completed the interactive safety training on the *Look, Stop and Fix* prototype. The interviewees represented different job titles, with nine being panel operators, six winch operators, five rock drill operators and three helpers.

The interviews were conducted by the researcher, assisted by two clerks from the mine, who acted as translators when required. Each interviewee's answers were written down by the researcher. Analysis of the responses yielded the following qualitative results.

1. Are you confident that you can do the job you were trained for? Are there any areas of your job you do not feel confident about?

The employees all seemed confident about doing their jobs, since 21 responded in the affirmative and two stated that they were "Very confident". Only two specified problem areas, namely "Equipment too heavy" and "We need more encouragement".

2. Do you think team work is important in your job environment? Why/why not?

All 23 interviewees responded positively, with some of the following being responses to the open-ended part of the question (each response is linked to the specific interviewee in square brackets, e.g. I1 refers to Interviewee number 1):

- "Share ideas and tasks" [I3]
- "One person might notice problem before the other" [I7]
- "Others may spot hazards" [I13]
- "People can look after each other" [I22]
- "Working as a team makes the job to be completed in time" [I23].

3. Why are accidents taking place if training is sufficient?

A variety of responses were received:

- Not complying with procedures [I1], [I3] and [I7]
- Using worn-out tools and equipment [I5], [I20] and [I22] ([I20] and [I22] stated they were forced to do this)
- Poor communication [I4] and [I19]

- Being ignorant and casual, not searching for hazards [I2], [I6], [I11], [I13], [I18] and [I23]
- Do not know or cannot explain [I8], [I9], [I12], [I15], [I16] and [I17]
- It is a dangerous job, accidents happen [I10] and [I14].

4. Can you remember the computer safety-training program you attended? What is the most important thing you remember about it?

All the interviewees remembered doing the training session. Most specified searching, seeing or identifying hazards as the most important issue, while others remembered the accidents that occurred in the simulation. Some also mentioned the feature of listening to the audio in their own vernacular.

5. Do you believe that this training helped you to work more safely? Why do you believe this?

Once again all the interviewees agreed that the training assisted them in working more safely in the real world. Some of the reasons given, were:

- Being more aware of hazards [I2], [I4], [I5], [I9], [I13], [I14], [I15], [I18] and [I22], e.g. "I'm more aware of the hazards and eliminate them as soon as possible"
- Learned to identify hazards [I3], [I10] and [I23]
- No one has been injured in any way since doing the training [I11] and [I20], e.g. "When I work I remember the things I saw on the PC"
- It shows accidents and causes [I1], [I7], [I12] and [I17], e.g. "It taught me that you should search for hazards immediately when you enter the cage right up until at your working place".

6. Does this form of training, the use of computers and simulations, work better than normal training methods? If it is so, why would you say it is?

Eight of the employees answered "No" or "Not necessarily", and indicated that hands-on practical training is still their preferred method. The other fifteen answered "Yes", mainly because it showed hazards and the accidents that can happen. One participant mentioned that the specific training program was a "wake-up call".

7. Would you think that other employees should also undergo this form of safety training (by using computers)? Why do you think so?

All the interviewees answered positively, with some of the reasons being:

- "We learn a lot" [I3], [I4], [I11], [I12] and [I13]
- "It makes you more safety conscious" [I2] and [I21]
- "You can see how accidents happen" [I1], [I7] and [I15]
- "It is easy to learn" [I8]
- "An eye opener" [I22].

8. What did you think of the program itself – was it realistic and could you easily recognise the hazards?

All the interviewees agreed that it was adequately realistic and some commented that the hazards were easy to spot.

9. If you could change the program – what would you change and why?

All interviewees responded "Nothing" and some indicated that everything was good, e.g. "Nothing, because whoever initiated this training will decrease the risk of accidents" [I8].

It became clear from the interviews that the training was well-accepted and perceived by the trainees to be valuable and necessary. The system made trainees more aware of hazards and they remembered the simulated accidents, as well as their causes. The interactive computer-based simulation methodology was deemed realistic and fit for purpose. Many indicated that they preferred that type of training to classroom-based training methods, and believed that the VR training was contributing to a safer work environment and increased hazard awareness.

6.8. Conclusion

This chapter has argued that the usability and usage of a system or product depend on its context of use. Context analysis is therefore an essential pre-requisite for any work on usability. The results of a usability context analysis provide useful input to the processes of specifying usability requirements, constructing a design prototype and evaluating the prototype with typical end-users.

Section 6.2 provided a background to UCA, while Section 6.3 presented UCA examples from the literature. Section 6.4 reported the application of the UCA conducted at the platinum mine, and the findings were discussed in Section 6.5. Sections 6.6 and 6.7 gave more details on the *LSF* prototype and reported on the findings of an early evaluation.

Based on observations during the evaluations described in Section 6.7 and the feedback received from participants, the following can be concluded:

- Miners require thorough pre-training on computer literacy before using the virtual reality simulation. The amount of time required for this seems to be more for the miners with lower levels of school education. The software that was developed for the pre-training worked well, but some miners required personal assistance.
- All the miners experienced the system as contributing positively towards safety awareness, hazard recognition and information on correct procedures in dealing with hazards.
- A number of improvements can be made to the system, including more exercises in the pre-training software to better prepare users for the controls used in the simulation and adding an empty mine to the pre-training, which users must walk through before attempting the hazard simulation system. This will ensure that the miners become more comfortable with the controls of the system.
- The system should be customised according to the safety standards at a particular mine, since the standards are not identical at all the mines.
- Virtual reality can be further utilised to address job-specific training needs.

The introduction of the interactive simulation training system has been beneficial to the workforce. No statistical data is yet available to indicate whether the accident rate has been reduced. That would require extensive use of such systems and several years of data collection and analysis. The *LSF* prototype system was well received and provided an interesting and enjoyable alternative to other training programs, while at the same time it also improved the safety culture and awareness of the workforce.

The findings of this early evaluation led to refinements to the *LSF* prototype, which resulted in an improved version of *LSF*, detailed in Section 5.6.7. Due to the success of the improved version, the Mine requested that future improvements should focus less on generic hazards, but more on the causes of actual incidents at the Mine. To determine these causes, a case study was undertaken to analyse recent incident data.

This chapter addressed Research Subquestion 3, by applying a UCA to determine the contextual requirements for virtual reality training systems for the mining industry. These requirements are described in Section 6.5, but are also augmented by the findings of the case study in Chapter Seven.

The deliverable of the chapter is the *LSF* prototype system. This prototype simulates the underground working areas, incorporating potential hazards. Mine workers need to spot the hazards, identify them correctly and indicate appropriate actions to be followed in response to each hazard. *LSF* also shows animations of the possible results of ignoring or not correctly addressing the hazards.

The evaluation of the *LSF* prototype, as discussed in Section 6.7, was conducted with instruments designed more by pragmatism than by application of theory. This evaluation served effectively to investigate the first version of the *LSF* prototype, which led to an improved version, but also played an important role in indicating the need for a formal set of evaluation criteria. Chapter Seven covers the evaluation of both the improved *LSF* prototype and the *ISGC* prototype, using formal criteria.

In this chapter, the first two of the four DBR cycles, referred to in Chapter Five, were discussed in detail. Chapter Seven provides insight into the next cycle, including details of the case study that led to the design and development of the *ISGC* prototype.

Chapter Seven
Prototype Design

7.1. Introduction

After successful deployment of the *LSF* prototype (as discussed in Chapter Six), management at the Mine requested that the training be extended to focus not only on generic hazards, but more specifically on the major causes of accidents. Due to falls of ground being the largest contributor to mining injuries, a case study was done at a selected site, namely Impala Platinum Mine, to analyse the circumstances relating to fall-of-ground incidents. Just as a usability context analysis preceded the design of the first prototype, *Look, Stop and Fix (LSF)*, similarly the results of the case study informed the design of a second interactive desktop VR prototype, the *Interactive Simulated Geological Conditions (ISGC)* prototype.

This chapter presents the details of the case study, which was an input to Cycle 3, as well as the design of the *ISGC* prototype developed and used in the third cycle of the DBR model, as described in Section 5.6. The design and development of *LSF*, which was used in the first and second cycles of the DBR model, was discussed in Section 6.6, whereas the *ISGC* prototype was used in the third cycle.

The *LSF* prototype was designed in collaboration with mine training practitioners through an iterative process of continuous design refinements, but no formal design methodology was followed. By contrast, for the design of the *ISGC* prototype, the researcher aimed to apply information systems theory and to base the design on an established design model. However, due to the novelty and distinct features required in interactive VR training systems for the mining industry, the researcher was not able to find an established design model that was directly relevant to this context. The established *interaction design lifecycle model* of Rogers, Sharp and Preece (2011) was therefore adapted by the researcher for the design of VR systems, and this extended interaction design lifecycle model was then applied to design *ISGC*.

Figure 7.1 shows the layout of this chapter. Section 7.2 discusses the case study analysing the details relating to falls of ground at a large platinum mine, which led to the design of the *ISGC* prototype. Section 7.3 explains the design lifecycle model used for the design of *ISGC,* namely the *extended interaction design lifecycle model* proposed for VR training systems. This is followed by information on the detailed design and development of *ISGC* in Section 7.4. The chapter concludes in Section 7.5 with design improvements made to the prototype after a number of evaluations.

The sections in colour in Figure 7.2 indicate which part of the research process is covered by this chapter, that is, the case study as input to Cycle 3, and the analysis, design, development and improvement of the *ISGC* prototype. *ISGC* was eventually deployed as a real-world practical solution at several mine training centres.

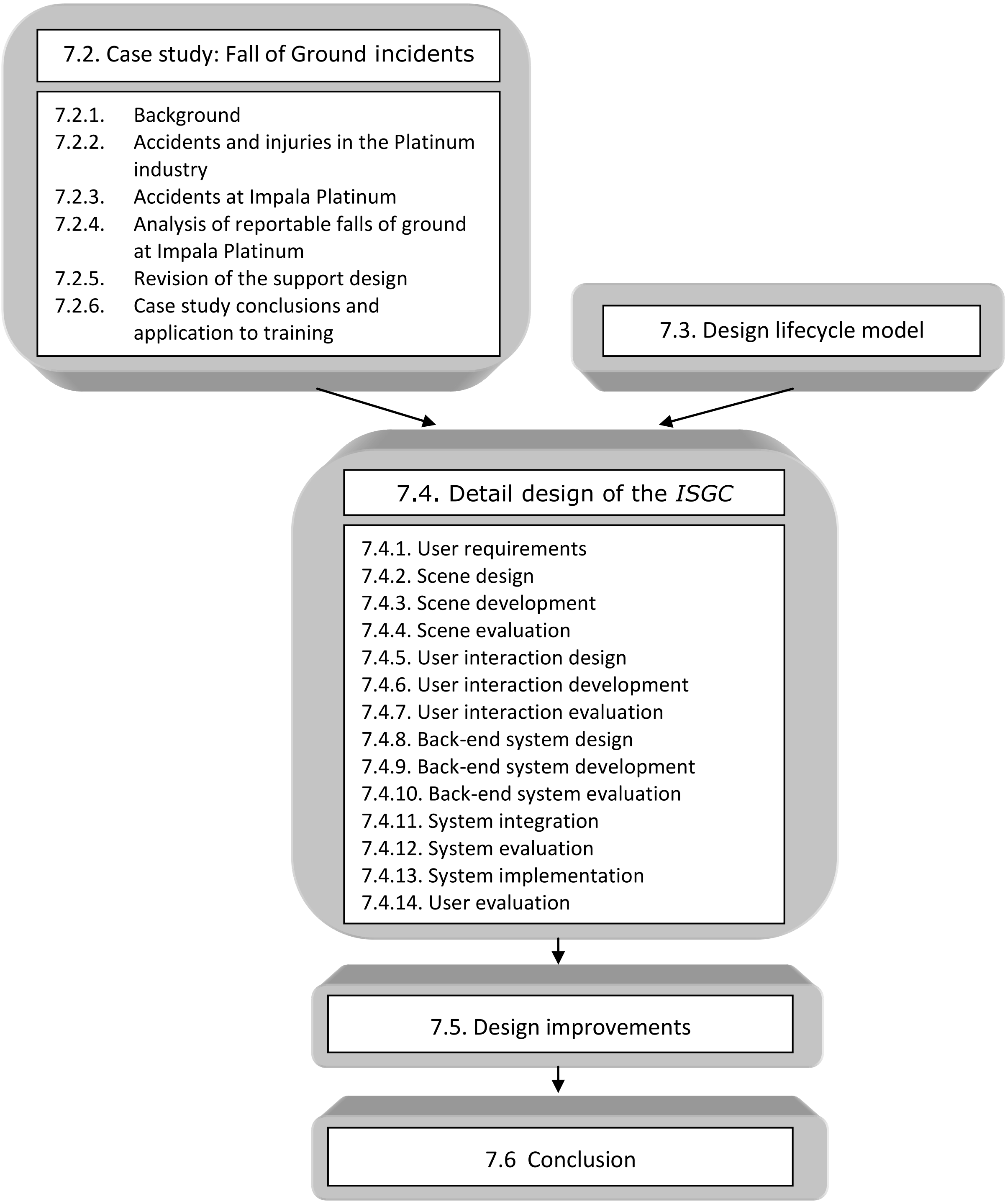


Figure 7.1: Layout of Chapter Seven.

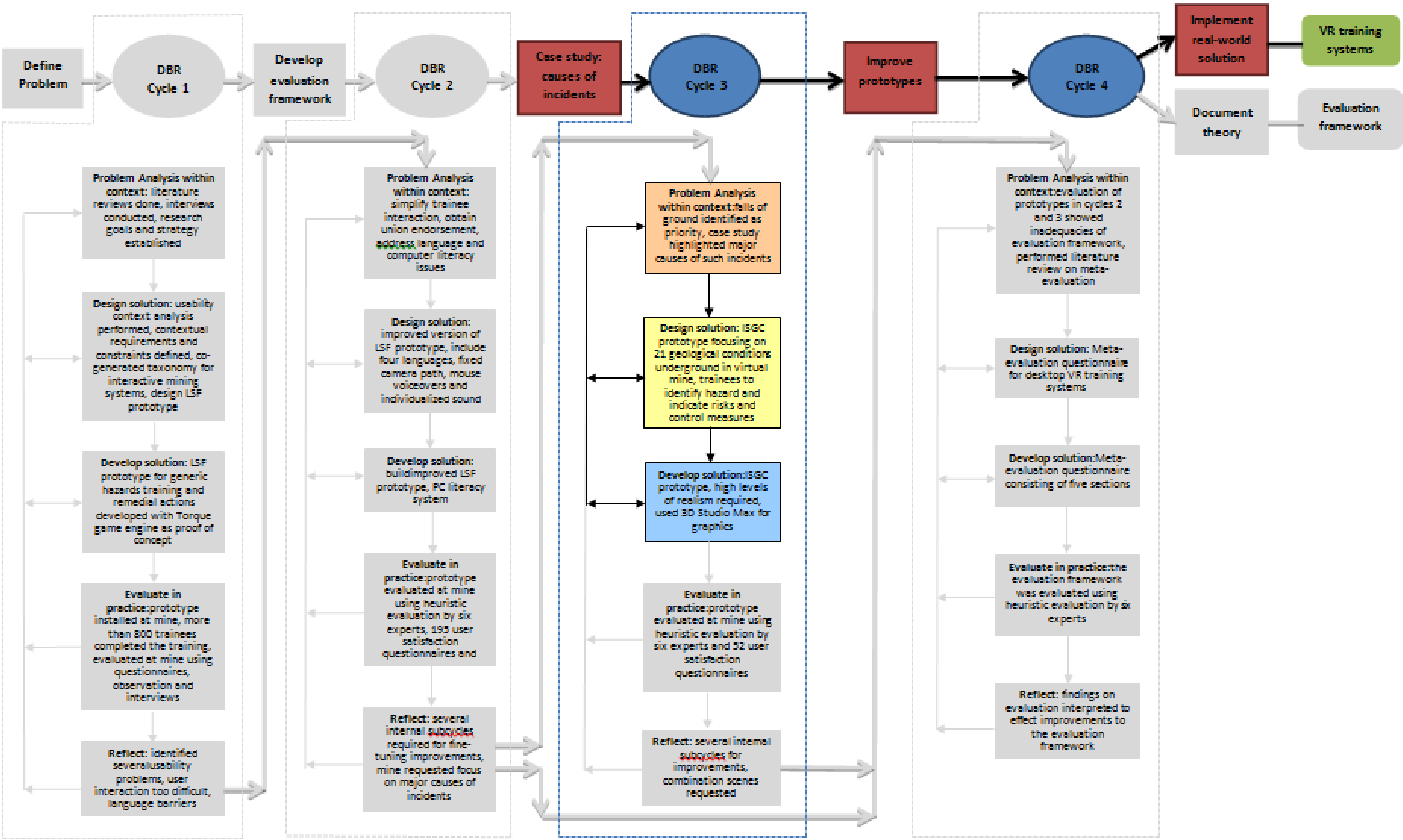


Figure 7.2: Research process flow discussed in Chapter Seven.

7.2. Case study: fall-of-ground incidents

Due to falls of ground being the greatest contributing factor to mining injuries, an in-depth investigation of this cause of incidents was warranted. In the case study, details relating to falls of ground at Impala Platinum, one of the world's largest platinum mines, were analysed. As indicated in the timeline in Section 5.6.8, this case study was conducted in 2010, and was based on data that had been collected during the period 2003–2009.

The case study was done to identify learning content and to inform the design of *ISGC*. The next five subsections, 7.2.1 to 7.2.5, present details of the case study, and the relevance of this information to the design of ISGC is discussed in Subsection 7.2.6. The contextual requirements discussed in this section also supplement the discussion of Research Subquestion 3, as presented in Section 6.5.

RQ3	What are the contextual requirements for virtual reality training systems for the mining industry?

7.2.1. Background

Impala Platinum Holdings Limited (Implats) is one of the world's most significant producers of platinum group metals (PGMs), which comprise platinum, iridium, palladium, osmium, ruthenium and rhodium. During 2007 South Africa produced thirteen million ounces of PGMs, making South Africa the world leader in PGM production. Implats produces approximately 25% of global platinum output (Implats, 2010a).

The catalytic properties of PGMs make them ideally suited for countering the effects of air pollution within the automotive industry. About 50% of the world's cars and more than 90% of new cars are fitted with platinum catalytic converters. Platinum is also used for jewellery and, in the petroleum industry, it is used to upgrade the octane content of gasoline.

The PGM orebody is a thin reef covering an extensive area, which makes the mining of PGMs similar to the methods used in gold mining: the reef is drilled and blasted to advance the face, after which support is installed for control of the hanging wall. These control methods, called strata control, differ significantly from gold mining in that gold reefs are sedimentary deposits and platinum reefs are igneous rocks (CoM, 2010b). The strata control systems used, are extremely important in order to prevent falls of ground during the mining process.

Implats employs approximately 46 500 people in South Africa and 6 500 in Zimbabwe, including large numbers of contractors. The Implats group has various mines operating in the two most significant known PGM orebodies in the world: the Bushveld Complex in South Africa and the Great Dyke in Zimbabwe, one of them being Impala, situated on the western limb of the Bushveld Complex, near Rustenburg in the North West Province.

During 2009, Impala Platinum was responsible for 56% of Implats' total production and approximately three quarters of net profit. Currently Impala's mining operations comprise 14 shafts, with 3 new shafts being developed. Approximately 20 000 employees work underground at Impala. The average mining depth is contantly increasing and is currently approaching 900 metres, with the deepest workings at approximately 1200 metres. About 17 million tons of ore are mined annually, yielding approximately 1.2 million ounces of platinum, plus additional quantities of other associated PGMs.

The mining process involves teams of seven to fifteen persons that drill, blast, clean and support panels at stoping widths of 1.0 m. The blasted rock is removed by scrapers and support is installed by hand. Two reefs are extracted, namely Merensky (named after Hans Merensky who first discovered platinum in the Bushveld Complex in 1924) and Upper Group 2, known as UG2 (Implats, 2010c).

7.2.2. Accidents and injuries in the Platinum industry

Figure 7.3 depicts the fatality frequency rate (FFR) from 2003 to 2009 in the platinum industry due to all causes. These figures are plotted against the 2013 milestones that were discussed in Section 4.3.4. The FFR is measured as the number of fatalities per million hours worked. The graph displays the FFR of the four largest platinum mining companies in South Africa: Implats, Anglo Platinum, Lonmin and Northam Platinum, with the red line indicating the average of the platinum industry and the black line the annual targets that should have been achieved on the way to the 2013 milestones. In comparison to other mines, the graph indicates that Implats' fatality frequency rate increased from the lowest during 2004–2006 to the highest in 2009. The company's position thus changed from the best performer to the worst performer. These statistics led to a re-evaluation of the strata control systems in Implats.

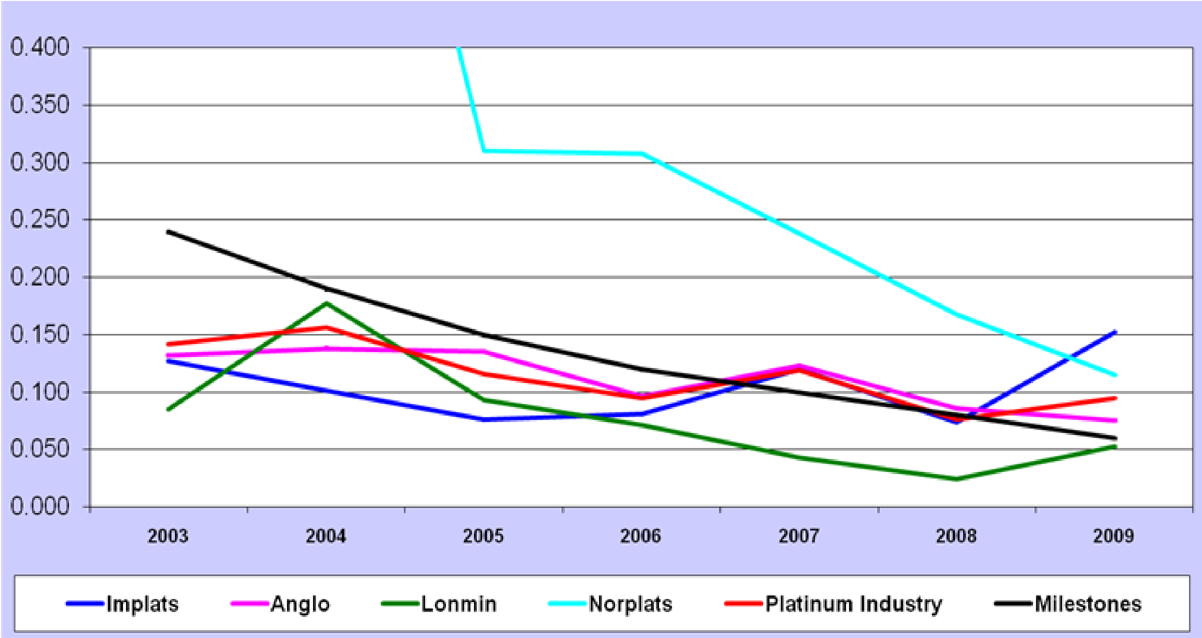


Figure 7.3: Platinum Industry Fatality Frequency Rate towards the 2013 milestones.

7.2.3. Accidents at Impala Platinum

The next three graphs indicate the rock-related accident trends at Implats. Figure 7.4 displays the number of fatalities at Implats caused by falls of ground (FOGs) in comparison to fatal accidents caused by other factors. This information is shown for each financial year from 2003 to 2009, where the financial year runs from July to June. It is clear from Figure 7.4 that, except for 2006 and 2009, FOGs remain the main cause of fatalities at the Implats mines, as is also the case in the South African mining industry in general. The data for 2010 is even worse as Impala suffered their worst ever FOG accident with 9 people killed in July 2009 when they were trapped by falling rocks approximately 1000 metres underground at Impala's No. 14 Shaft (Lourens, 2009).

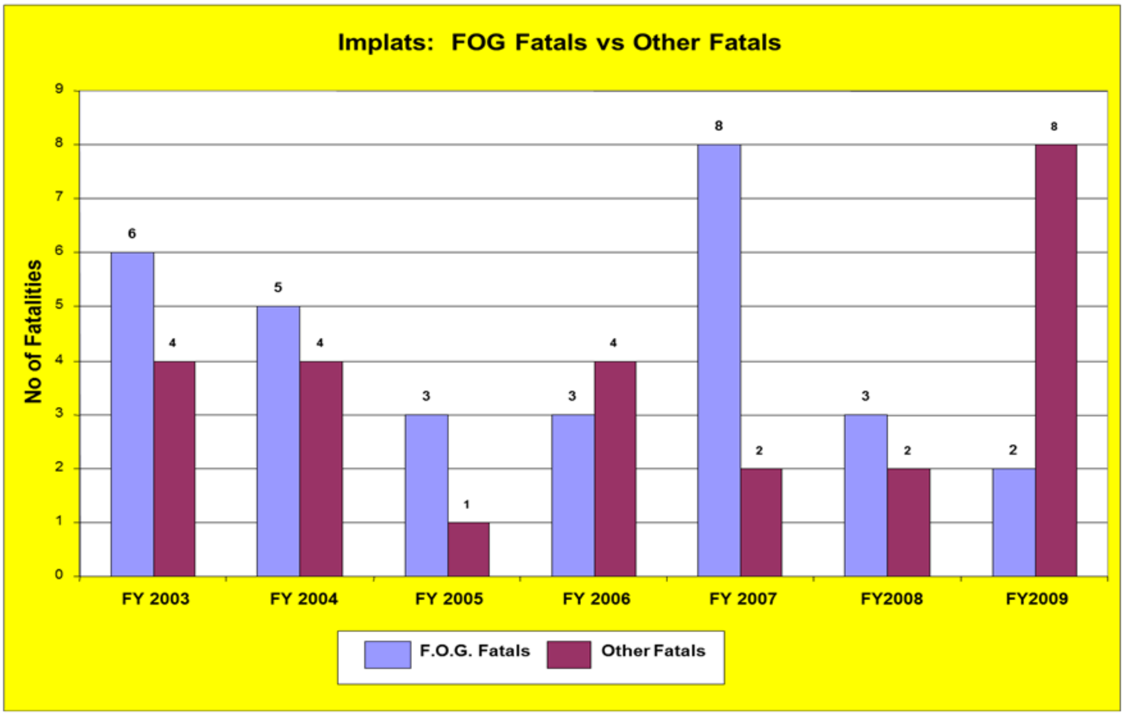


Figure 7.4: Fall of Ground fatalities compared to other fatalities at Implats.

Figure 7.5 indicates an upward trend in the number of reportable rock-related incidents. Reportable incidents are falls of ground larger than 10 m^2 in area or 5m^3 in size. All such incidents must be reported to the Department of Mineral Resources.

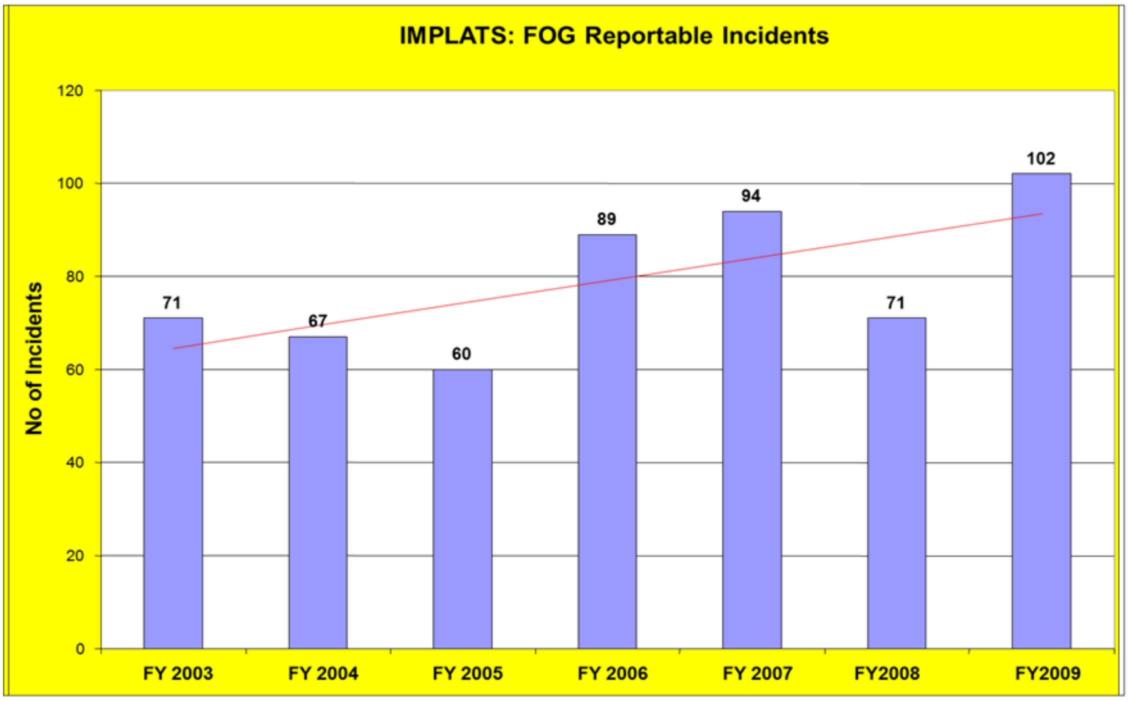


Figure 7.5: Reportable incidents at Implats from 2003 to 2009.

Figure 7.6 shows the number of injuries for each financial year. Although the number of accidents increased, a downward trend in injuries is noticeable since 2006.

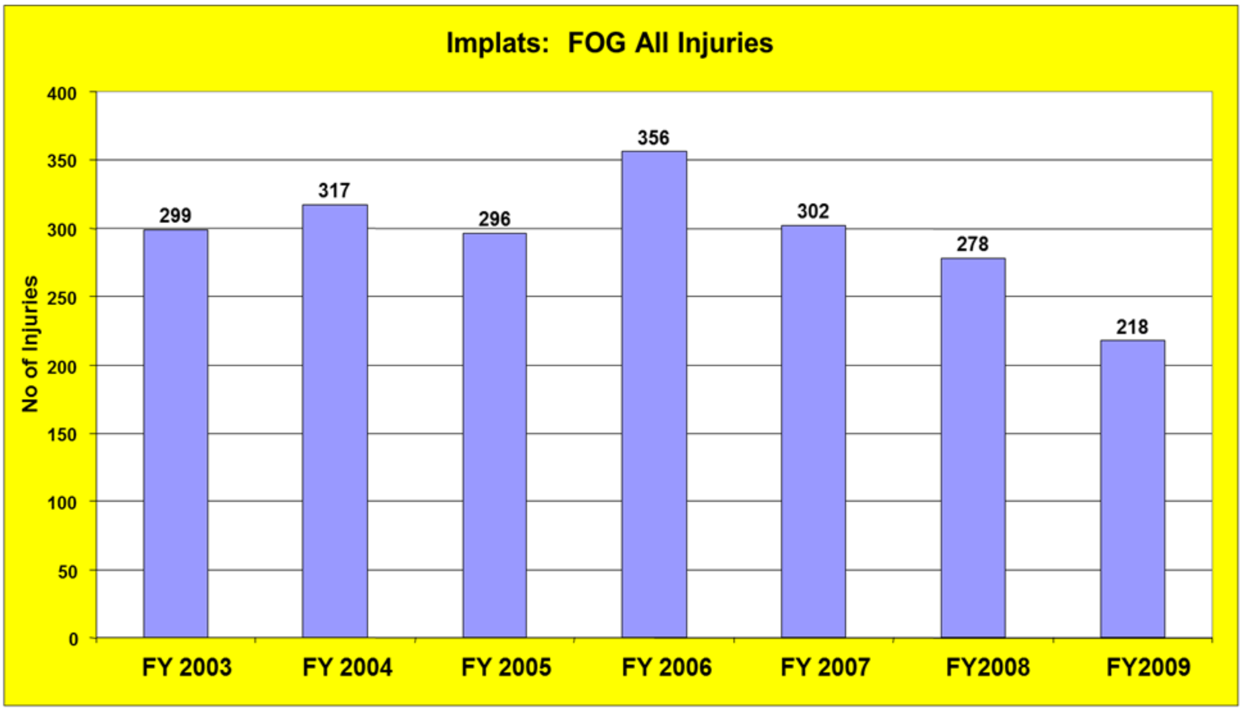


Figure 7.6: Number of injuries at Implats due to falls of ground.

Due to the increasing number of reportable FOGs and fatalities due to FOGs, the Rock Engineering Department at Implats decided in 2007 to implement three additional safety measures:

- in-stope roof bolting on the Merensky reefs,
- increased roof bolting on the advance strike gullies (ASGs), and
- use of rock netting, also called aerial support, on the UG2 reef.

Figure 7.7 shows an example of how rock netting and roofbolts can be used to support the hanging wall.

Figure 7.7: Rock netting installed at Impala No. 10 shaft.

The following are some of the advantages of introducing the in-stope roof bolting support system with aerial support such as nets:

- The permanent support distance to the face can be reduced.
- The number of smaller falls of ground occurring near the face can be reduced.
- The use of temporary mechanical props that must be installed and removed daily, can be minimised.
- The use of roofbolts reduced the after-the-blast span from permanent support (mine poles) to the face from 4,0 metres to 1,5 metres, making it safer for cleaning operations.
- In terms of a complete stope, the unsupported area was reduced from 100 m² (25 m x 4 m) to 37.5 m² (25 m x 1.5 m) after the blast. This is an improvement of 260% on a 25 m long panel.

As can be seen in Figures 7.4 and 7.5, these additional measures improved the fatality rate, but were not successful in decreasing the number of reportable FOG accidents. This prompted a decision by mine management to strengthen the training program. Following the success of the *LSF* prototype, a request was made for a VR training system focusing on FOG hazards. This led to the development of the *ISGC* prototype, which is addressed further in Sections 7.2.6 and 7.3.

7.2.4. Analysis of reportable falls of ground at Impala Platinum

During the period July 2009 to April 2010, a total of 46 reportable FOGs occurred at Impala Platinum in Rustenburg. Of these, 26 collapses occurred while mining the UG2 reef and 20 at Merensky reef extractions.

With the aim of preventing repeat occurrences of these FOGs, wherever possible, and to avoid injuries to workers entering hazardous areas, it was required that *ISGC* should include simulations of the conditions that caused these FOGs. In order to decide which geological conditions to include in the design of ISGC, this section reports on the details of these 46 FOGs at Impala Platinum. Furthermore, to enable them to build realistic simulations of these hazards, the developers needed to know where the geological conditions were situated in relation to the face wall, which geological hazards were present at the FOG areas, where the falls of ground occurred, which geological features were involved, and the sizes of the rocks that fell. This information is presented in the following five subsections.

7.2.4.1. Location of the falls of ground

Impala Platinum employs conventional mining methods at all its shafts, but also uses mechanised mining methods at No. 12 and No. 14 shafts. Mechanised mining using trackless mining moving machinery, also known as TM3, accounts for approximately 13% of Impala’s production (Implats, 2010b).

As indicated in red in Figure 7.8, a high percentage of these FOGs occurred at the mining face area. In seven cases, which represents 16% of the conventional mining operations, a total collapse of the face occurred and four accidents (9%) also occurred in TM3 operations. No accidents occurred in the ASG face area or in sidings, and 31% of the accidents occurred in the back areas of stopes.

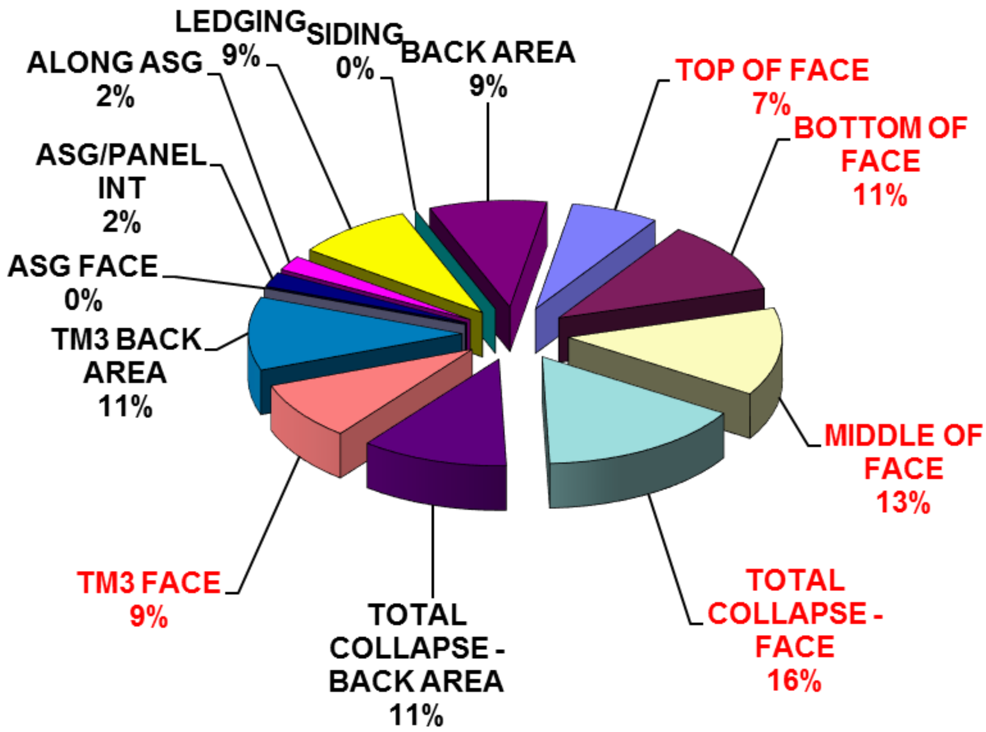


Figure 7.8: Location of FOGs July 2009 to April 2010 at Impala Platinum.

7.2.4.2. Hazards involved in the falls of ground

Figure 7.9 indicates the geological hazards present at the FOG areas. It clearly shows the dangers of working in an unsupported face area or in face areas after blasting (indicated in red). Other notable geological hazards include low angle joints, shear zones, dykes and faults, while off-reef mining also contributed to 12% of the collapses. A further alarming finding was with regard to support compliance. Investigations revealed that in 70% of the FOG cases, the support was sub-standard.

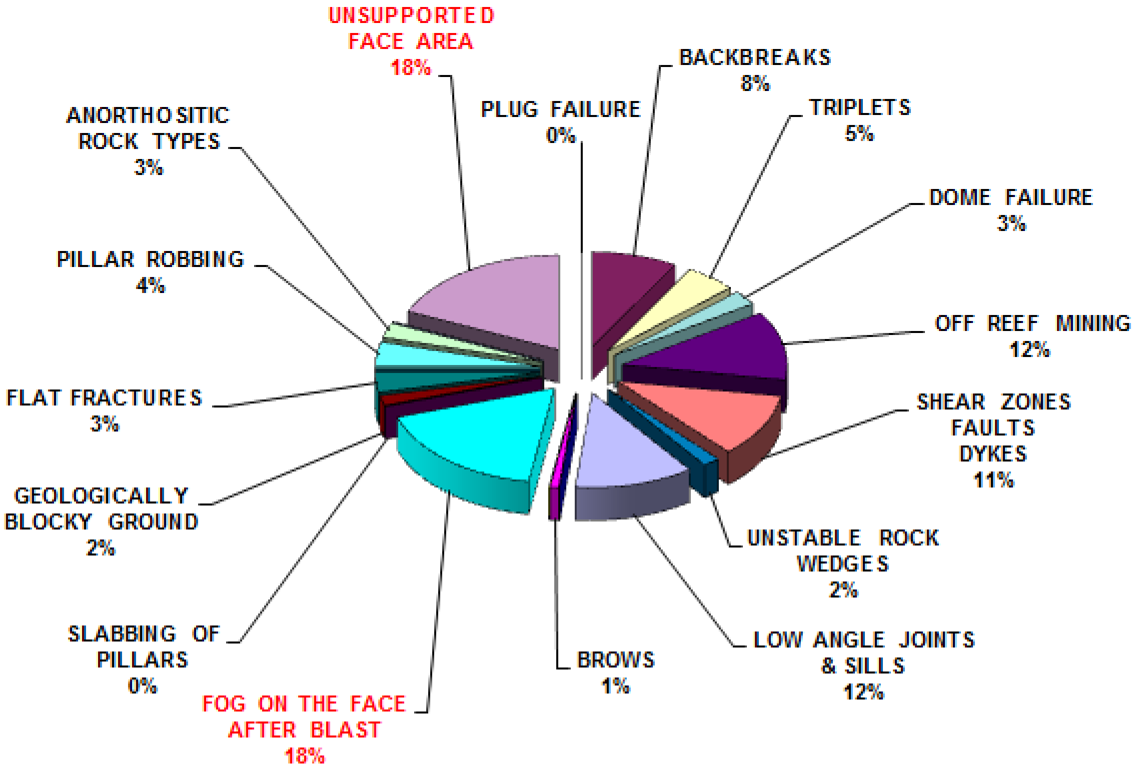


Figure 7.9: Hazards involved in the FOG's July 2009 to April 2010 at Impala Platinum.

7.2.4.3. Distance of falls of ground from the face

The number of FOGs occurring in proximity of the face area, are indicated in Figure 7.10. The majority of the falls occurred within four metres of the face, once again highlighting a need for improved support close to the face. The high number of falls further than fifteen metres from the face is also a factor of concern, given that those FOGs occurred in areas with permanent support.

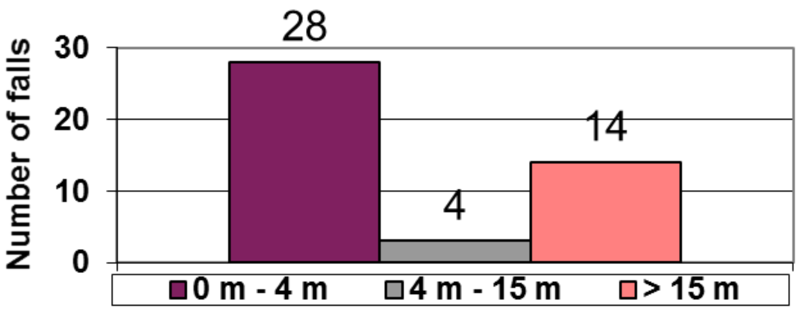


Figure 7.10: Number of falls at specific distances from the mining face.

7.2.4.4. Geological features involved in the falls of ground

As indicated in red, it is clear from Figure 7.11 that jointing caused the most problems, since joints and low-angled joints were present in 31% of the collapses. Other major contributors were pegmatite veins, faults, dykes and potholes.

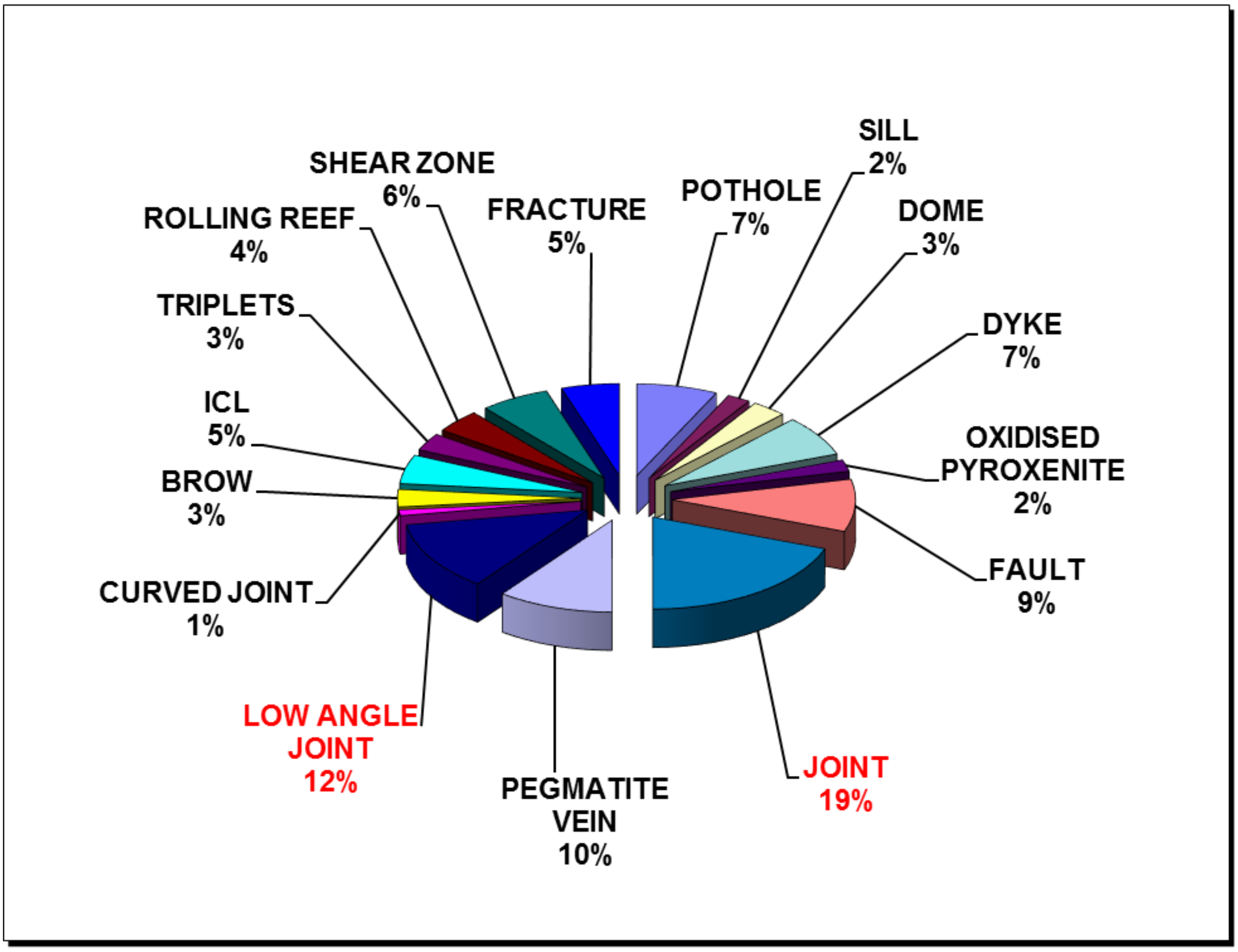


Figure 7.11: Geological features involved in the FOG's.

7.2.4.5. Falls of ground rock dimensions

In terms of the size of the rocks that fell, the average rock dimensions for the 46 measured incidents were 12 m in length, 7 m in width and 1.1 m thickness, as indicated in Figure 7.12.

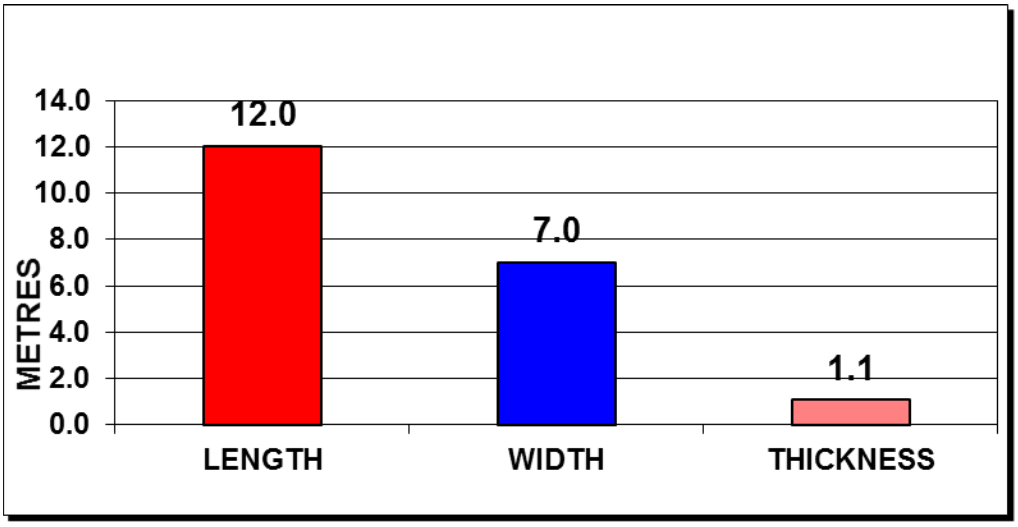


Figure 7.12: Average rock dimensions of the 46 FOG's.

7.2.5. Ground control districts

A Ground Control District (GCD) is an area of a mine where similar geological conditions exist. This gives rise to a unique set of identifiable rock-related hazards, for which a common set of strategies can be employed to minimise the risk resulting from mining. Ground control means the ability to predict and influence the behaviour of rock in a mining environment, having due regard for the safety of the work force and the required serviceability and design of the mine.

As a result of the number of major collapses occurring on Impala, the Rock Engineering Department, in conjunction with the mining personnel, started identifying areas with similar ground conditions in order to define ground control districts. The result of this exercise yielded nine GCDs at Impala, as indicated in Table 7.1.

Table 7.1: Ground Control District classification at Impala Platinum.

GCD code	Ground Control District Description	Major generic hazard description
A	Normal Ground	Low risk from geological features
B	Surface Protection	0 - 30 m No Mining 30 - 100 m Shallow mining restrictions
C	Curved Joints	Wedge type failure
D	Coarse pyroxenite Spotted anorthosite	Large flat failures, extension fractures
E	Rolling Reef	Associated with curved joints, domes & various reef types on the same panel
F	Blocky ground	Associated with extensive jointing, faulting, shear zones, etc. on the various reef horizons
G	Triplets or ICL < 0.3 m above UG2 Reef Contact	Narrow beam can result in falls of ground between support units.
H	Low angle joints on UG2	Series of domes that intersect into the triplets and results in major falls of ground
S	Seismicity	Seismic risk associated mostly with crush type events (pillar or strain bursts) and less with slip type events (geological features)

The GCDs were colour-coded relating to the relevant support strategy. GCDs were defined based on the following:

- The presence of dominant or persistent geological features.
- The potential fall-out height for the face area.
- The potential fall-out height for the back area.
- Risk factors associated with different areas or phenomena.

A Ground Control District Plan displays all the applicable ground control districts of a mine to a scale of 1:2500. Figure 7.13 is an example of a GCD Plan, showing No. 10 Shaft at Impala, colour-coded according to its GCDs.

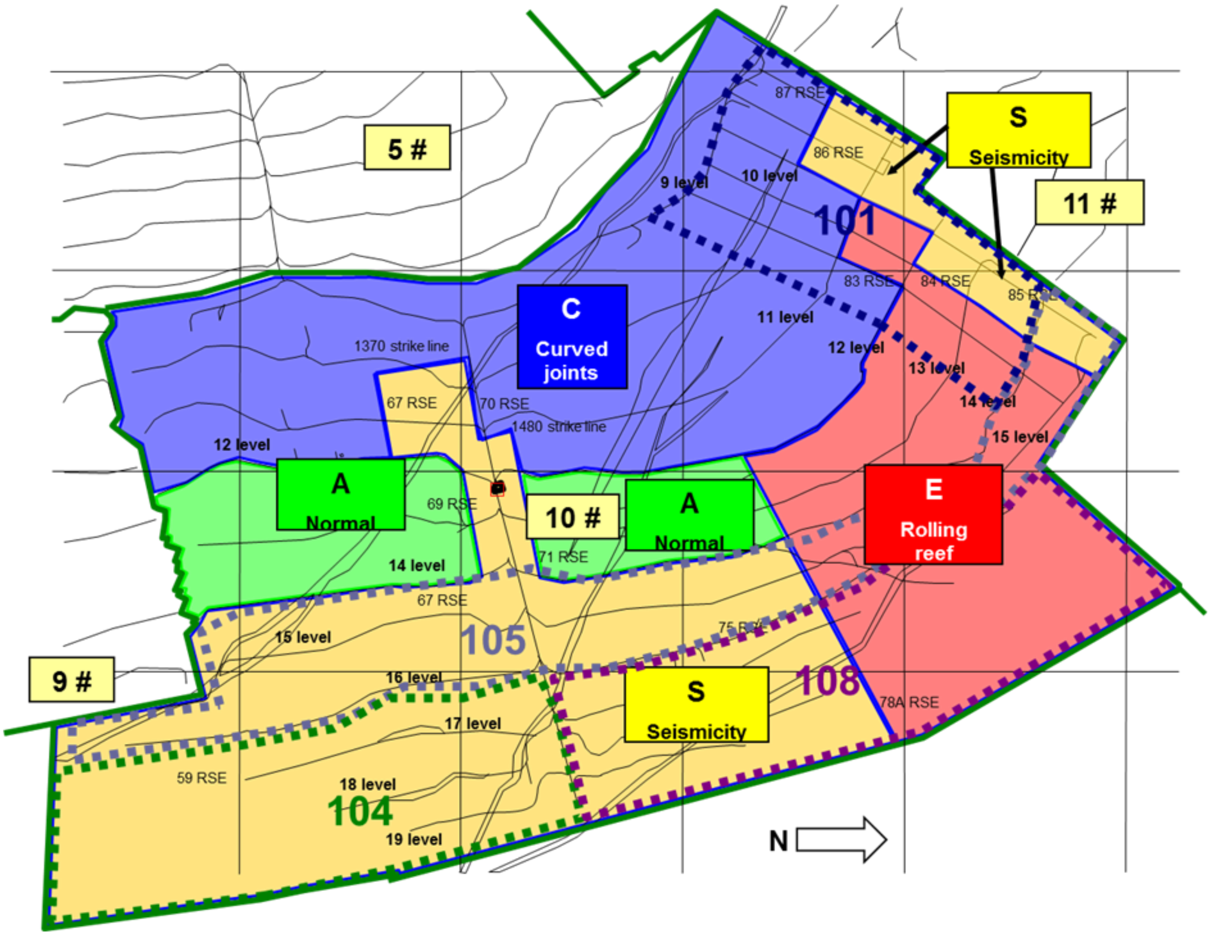


Figure 7.13: Ground Control Districts and Support Strategies for No. 10 Shaft.

Developing an effective GCD Plan requires a formal method of gathering information regarding FOGs, keeping the database updated regularly, and plotting geological features onto plans. All FOGs should be analysed, not only the large collapses and those causing accidents.

Although the GCD determines the support system to be implemented, the support strategy does not cater for all conditions within a declared GCD area. For example, within GCD "A" the panel might intersect a fault associated with jointing, therefore different support strategies should be considered to mitigate the hazards along the fault. Multiple combinations of different geological conditions can occur within an area. It is also evident that, in some instances, areas need to be barricaded off and abandoned, since no support will suffice.

During an interview with Mr Noel Fernandez, Chief Group Rock Engineer at Implats, he pointed out that the GCD support strategy does not replace good judgement on the face. Good judgement is enhanced through training, experience and the ability to make safety decisions without being influenced by production targets and production bonuses. He emphasised the fact that safety training on geological conditions at the mine must be revised and improved (Fernandez, 2010).

7.2.6. Case study conclusions and application to training

During the case study it became clear that Implats had a major problem relating to the occurrence of FOG incidents. Major strides were required to achieve safety targets. Another important aspect was that, once a mine achieves its annual target, the challenge lies in consistently remaining below the milestones.

From the analysis of the 46 FOGs the following major conclusions emerged:

- Most FOGs occurred on UG2 panel faces.
- The falls frequently occurred after blasts in areas normally unsupported for 4 m from the face.
- Support compliance is a problem.
- In 98% of the cases, geological features were involved, with joints, pegmatite veins, faults, dykes and potholes being the main features involved.

The case study findings described in the previous five subsections, provided valuable information for the design and development of training systems on FOGs. Not only did the case study uncover which geological conditions where present in the analysed FOGs, but also where these FOGs were located, the geological hazards present at the FOG areas, the rock dimensions of the rocks that fell, and the importance of a GCD plan to indicate the support strategy for each GCD area. This information was used to design the content of the *ISGC* training system, as well as the computer-generated imagery to portray the simulated hazards and rock fall animations.

Of particular interest to this study, is the recommendation from the Chief Group Rock Engineer that training should be improved. Moreover, an emphasis was placed on the importance of training to foster good judgement on the part of trainees. The next sections discuss how these needs were addressed by the *ISGC* prototype, which is an interactive desktop VR training prototype designed to improve awareness, identification and support of geological hazards.

7.3. Design lifecycle model

This section addresses Research Subquestion 4 of this study:

RQ4	What is an appropriate design lifecycle model for interactive desktop virtual reality training systems?

In DBR Cycles 1 and 2, discussed in Chapter Six, the *LSF* prototype underwent a cyclic design, develop, implement and evaluation process, which led to the successful adoption of the system by mine management at the end of Cycle 2. Due to the fact that *LSF* only consisted of generic hazards, Implats requested that a further training system be developed, this time focusing on falls of ground and the associated geological conditions prevalent in recent FOG incidents at its Rustenburg mine. A case study was conducted to analyse the recent FOG incidents, and its findings are presented in Section 7.2. These findings informed the design and development of this new training system, called *Interactive Simulated Geological Conditions (ISGC)*. As indicated in Section 5.6, where the research process flow was discussed, the analysis, design and development of *ISGC* formed part of DBR Cycle 3, which is described in detail in Sections 7.3 to 7.5. A pragmatic, less formal approach was followed to design *LSF*, but it was decided to base the design of *ISGC* on a formal design model.

As mentioned in Section 7.1, the researcher had not been able to find an established design model directly relevant to this context. Chapter Three highlighted a number of user-centred design methods: user-centred system design, learner-centred design, interaction design and usability engineering. As mentioned in Section 3.2.1, an advantage of a user-centred approach is that it prevents projection of the designer's personal view, and focuses rather on the most pertinent stakeholders, such as the target users and organisational management. By contrast, designers working independently, without input from others, may personally find the system easy to use, yet may overlook critical design weaknesses that would pose problems for others.

As this study investigates the introduction of new technology and innovative training methods into the mining industry, it was important to include all stakeholders in the development of the proposed approaches in order to gain acceptance in the industry. This emphasises the role of a design approach that takes cognisance of multiple viewpoints and information from a variety of sources.

As discussed in Section 3.2.3, Rogers *et al.* (2011) advocate the term *interaction design* as an umbrella term covering interface design, user-centred design, web design, software design and interactive system design. The process of interaction design involves four main activities:

- The needs of the users should be determined, and the requirements established.
- Alternative designs should be generated, based on the requirements.
- Interactive versions of the design should be developed.
- The resulting products should be evaluated.

Figure 7.14 depicts the activities in the interaction design lifecycle model. This model has its roots in software engineering and HCI lifecycles, but Rogers *et al.* (2011) point out that it is not intended to be prescriptive, and is based on what they believe is practised in the field. More than one alternative design may be generated within this iterative cycle in parallel with others, or one alternative can be considered at a time in a serial approach. The final product emerges in an evolutionary fashion from the initial idea to a finished product which meets the prescribed usability criteria.

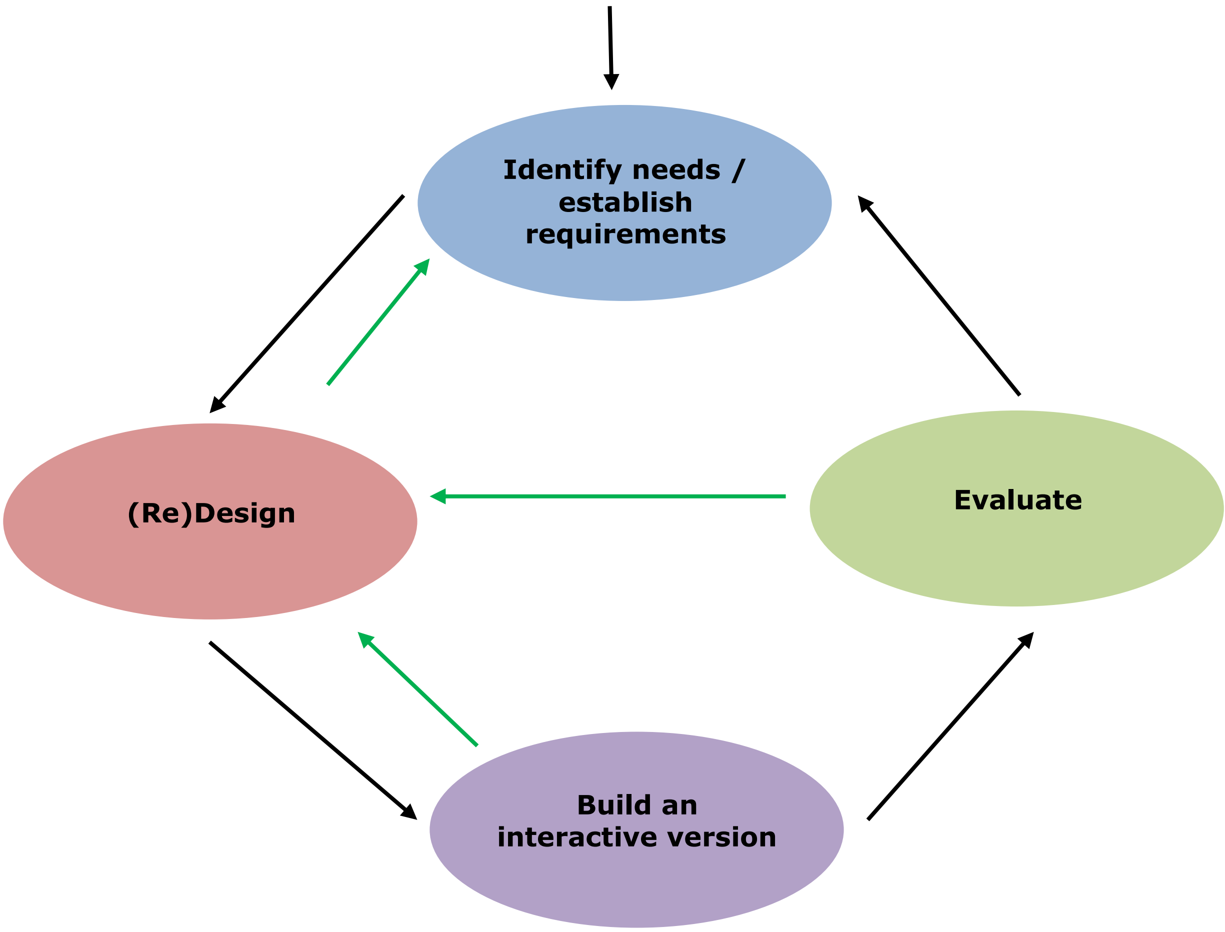


Figure 7.14: A simple interaction design lifecycle model (based on Rogers *et al.*, 2011).

Figure 7.14 also indicates the relationship between the activities in interaction design. The black arrows indicate the flow of activities, while the green arrows indicate iterative returns to previous activities. This may be due to:

- feedback from evaluations, where the developers need to return to the identification of needs or to the refinement of requirements, or they may move directly to redesign,
- generation of alternative designs in an attempt to address the identified needs and requirements, or
- the need to redesign a solution, which was not deemed feasible at the stage when the initial interactive version was built.

Interaction design is conducted by multidisciplinary teams of stakeholders. In this study, it proved difficult to involve the potential users, namely, the mine workers, in the early design stages for the following reasons:

- The nature of the work and high production targets result in mine workers being tired after their shifts and ready to go off duty.
- The users have had little or no previous computer exposure.
- Very few of them have had previous computer-based training, and would therefore have no points of reference in such discussions.

The users could therefore only be involved in the evaluation of developed, operational systems, and only during periods earmarked for training purposes. This necessitated working closely with the safety and training departments of the mines, where experienced officials had in-depth knowledge of the target users. In this process, safety practitioners and training officers participated in determining the needs and establishing the requirements of the envisaged *ISGC* safety training system, in evaluating alternative designs, and in phases of the actual development of the system.

For the design of the VR training prototypes used in this study, the interaction design lifecycle model of Rogers *et al.* (2011) was extended by the researcher to make provision for three simultaneous processes, and the subsequent integration thereof into a single product. Figure 7.15 depicts the design model used.

The project commenced with the identification of the user needs, after which various designs were developed. The green block contains the three design processes that run concurrently, namely:

- Scene design: design of the virtual environment, including the objects and animations required in the simulation.

- User interaction design: instructional design, system questions and answers, user feedback, system introduction and different language options.
- Back-end system design: design of the system workflow, programming and database requirements.

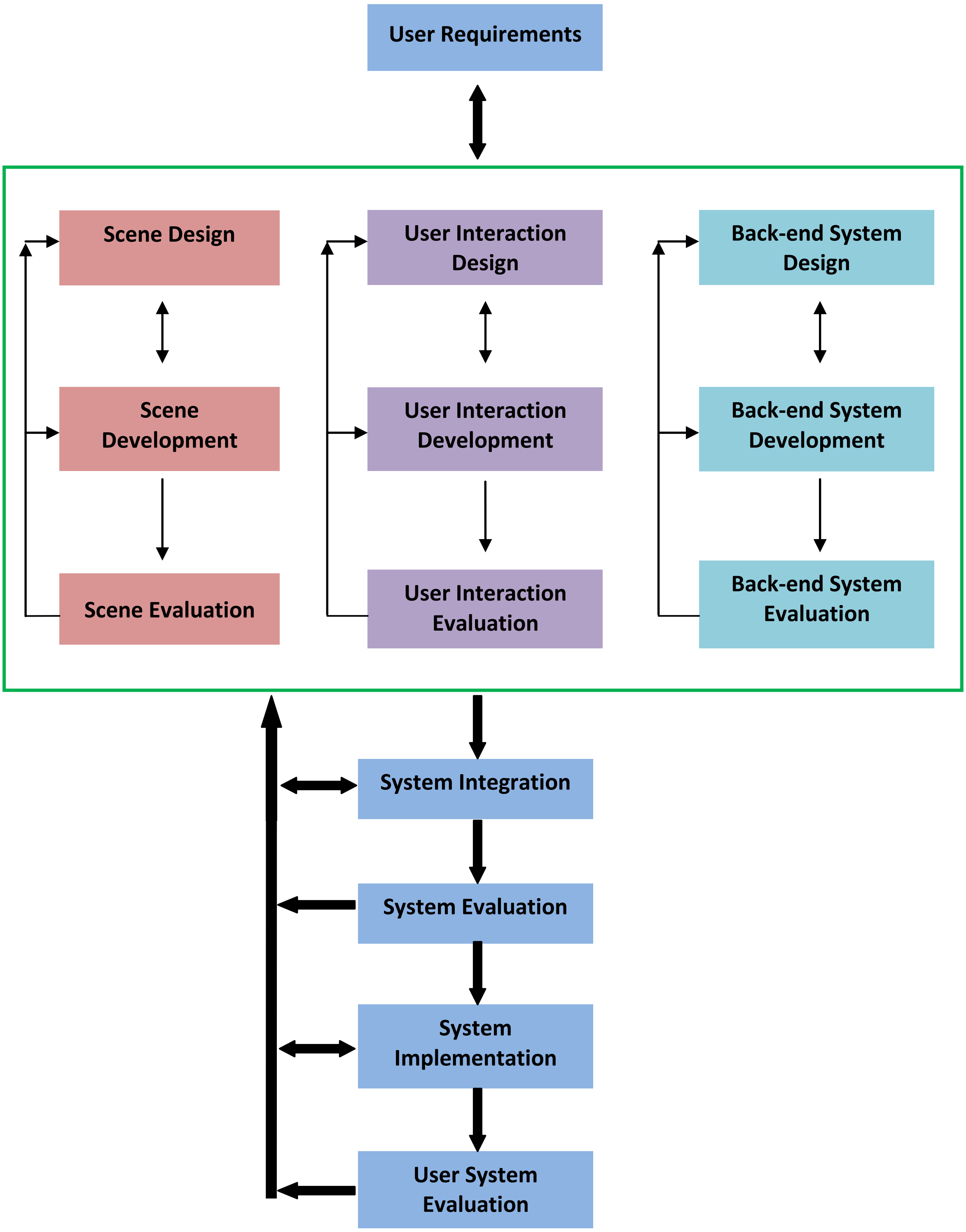


Figure 7.15: Extended interaction design lifecycle model (generated by the researcher).

Each design subcycle in the green block contains three activities: design, development and evaluation. The design activity includes both conceptual and physical design. The conceptual design describes what the system should do and what it should look like. The physical design considers the detail of the system, including the colours, textures, sounds, images, objects to be modelled and animations required. These details should be captured in a storyboard.

The scene design subcycle was repeated as least twice, since each scene in the prototype consisted of two parts: a hazard identification scene and an accident scene. In the scene development activity of the design cycle the storyboard was developed using applicable software. The developed products were then evaluated by the developers for quality control and also by the client. As indicated by the arrows, any activity can result in going back to improve a previous activity.

The user interaction design subcycle comprises the design, development and evaluation of the program flow and all user system interaction, including the text and narration, to be provided in different languages.

The back-end system design subcycle involves the selection of the appropriate development environment, including the programming language, database and reporting tool. This is then followed by the design, development and evaluation of the back-end system and database.

In line with Figure 7.15, once all the deliverables have been approved, they must be integrated into a training system. The training system is then evaluated to determine if all the deliverables of the design cycle work together as a unit, and satisfy the established requirements. If any discrepancies are found, improvements must be made in previous activities of the lifecycle, and evaluation must be performed again.

Successful evaluation leads to implementation of the training system at the client site. Once installed, evaluation by users can be done to ensure compliance to the stated requirements.

Even though the extended interaction design lifecycle indicates three separate design cycles which can run concurrently (scene design, user interaction design and back-end system design), in practice it could result in the same team being involved in the design and evaluation of all three activities. In terms of the development, however, different expertise will be required for modelling and animation, interface design, instructional design and programming.

It is therefore clear that the application of the extended interaction design lifecycle model leads to a system evolving through a process of iterative refinement, rather than simply being developed through a linear process. Two key concepts are formative evaluation and design iteration. Formative evaluation refers to the evaluation of design ideas and has the key aim to learn more about factors that impinge on design. Design iteration allows for the refinement and revisiting of any activity of the design. The following section discusses the design and development of the prototypes.

7.4. Detailed design of the *ISGC* prototype

Section 5.6 overviewed the application of four cycles of the DBR model, detailing the use of the *Look, Stop and Fix (LSF*) prototype for Cycle 1, an improved version of *LSF* for Cycle 2, and the specialised *Interactive Simulated Geological Conditions* (*ISGC)* prototype for Cycle 3, which is now addressed.

This section details the application of the extended interaction design lifecycle to the design of the *ISGC* prototype. Each of the activities depicted in Figure 7.15 is explained in the next fourteen subsections, in the following order: firstly the user requirements are defined; followed by the design, development and evaluation activities of the scenes, user interaction and back-end system respectively; and then the rest of the activities in Figure 7.15 from top to bottom, that is, system integration, system evaluation, system implementation and user system evaluation.

The system design team comprised a rock engineer, the mine training practitioner responsible for training in geological conditions, a mine safety training specialist, two VR system developers, and the present researcher. Other members of the rock engineering department at the Mine were also involved from time to time, especially during evaluation sessions.

7.4.1. User requirements

Due to falls of ground being the greatest contributing factor to mining injuries, Section 7.2 presented the results of a case study relating to causes and details of falls of ground. One of the outcomes of the case study was a recommendation for improved safety training. Following the success of the *LSF* prototype, which focused on generic hazards, mine stakeholders requested a training system with material that oriented learners on the different geological conditions prevalent in the mine's underground reef environment, in this case the Merensky reef. This led to the design, development and implementation of the *ISGC* prototype.

Based on the information collected during the case study, as well as the experience of the rock engineer and trainers, 21 geological conditions were selected as focus areas for the prototype. These conditions are briefly explained in Table 7.2.

Table 7.2: Geological conditions covered in the ISGC prototype.

Geological Feature	Short Description
Joint	A joint is a crack in the rock strata without any movement of the layers occurring. Joints disturb the continuity of the layering, especially the hanging wall and footwall layers surrounding the reef plane.
Fault	The layers of rock are not always one continuous layer but can be broken or discontinuous. A fault is a condition that occurs where the two sides of the discontinued rock strata move upwards or downwards in relation to one another.
Brow	A brow is a step in the hanging wall that remains after a piece of ground has fallen out of the hanging wall. The danger of a brow is that further falls of ground can occur if the brow is not adequately supported. Most brows are a direct result of poor drilling practices, where the hanging wall beam has been drilled into and broken.
Dyke	Dykes consist of lava that was forced in between rock layers and into cracks, faults and joints that did not bond with the surrounding rock. Dykes can be identified as rock layers that have a different appearance to that of the adjacent layers. The difference is in colour, structure or composition.
Prominent joint	A prominent joint is a joint that has infilling of 3 mm or more and contains water. When an excavation intersects water-bearing geological features, the water can lubricate joints, which can result in premature detachment of rock blocks, increasing the possibility of geological feature-bound falls of ground.
Dome	A dome is a curved joint which extends into the hanging wall. The main dangers associated with mining under domes are the low-angle fractures or joints that are very difficult to identify and are often overlooked or mistaken for a regular joint.
Reef in footwall	The reef rolls down to settle on a lower footwall layer.

Wedge	The intersection of geological features, such as joints, faults and dykes, results in the creation of wedges. Under certain conditions these wedges may be unstable, presenting a potential fall-of-ground hazard.
Stress fracturing	Stress fractures are commonly caused by stress exceeding the rock strength, causing the rock to lose cohesion along its weakest plane.
Shear zone	A shear zone is a zone of rock that has undergone deformation, and is surrounded by rocks with a lower state of finite strain.
Reef in hanging	Reef in the hanging wall occurs where the true reef and pyroxenite roll up into the hanging, creating unsafe conditions. This results in poor cohesion between layering contacts, especially in rolling reef mining.
Scaling	Scaling occurs when high stress levels on the reef result in unstable layers of rocks in the sidewall, which break away in an onion peel-like fashion.
Rolling reef	It is called rolling reef when the reef rolls down from a particular stratigraphic unit to settle on a different stratigraphic layer.
Pothole	A pothole is an occurrence of a bowl-shaped geological structure that causes a down-warping of the reef. Disrupted rock layering is found on the edges of potholes, along with small-scale faults and joints, resulting in the deterioration of ground conditions.
Pillar robbing	A pillar is a block of ore entirely surrounded by stoping, left intentionally for purposes of ground control. When the pillar width-to-height ratio is incorrect as a result of poor pillar cutting, scaling and slabbing of the pillar sidewalls can occur. Pillar robbing of one pillar causes adjacent pillars to be further loaded, which may exceed their design capabilities and has a negative effect on the stability of the surrounding area and working places.
Pegmatite vein	A pegmatite vein is a form of igneous rock consisting of extremely coarse granite resulting from the crystallisation of magma.
Off-reef mining	If the reef is not easily visible it can lead to off-reef mining, which reduces grades due to the dilution of ore with waste rock. Off-reef mining also increases the risk of seismicity.
Low angle joint	The dip of joints can vary from vertical to almost horizontal. In low angle joints the presence of infilling causes a lubrication of the joints that may result in movement and falls of ground.
Blocky ground	A key block is a rock mass of any particular size that internally supports or stabilises a number of adjacent rock masses in a workplace. Blocky ground has key blocks that fit into each other or fit together like a jigsaw puzzle.

Brittle rock	Brittle rock is rock that has a tendency to fracture under low stress without appreciable deformation.
Blasting Damage	Blasting damage is damage to the hanging wall due to sub-standard blasting practices, e.g. overcharging of holes or incorrect blast timing.

Rock in which and through which mining is done, is formed in layers, called strata. Each rock layer is formed from a different type of rock, each one with its own distinct properties and characteristics. Strata control is concerned with the support of fractured rock around underground excavations in order to achieve safe conditions.

The learning material in the *ISGC* prototype had to include visualisations to present means of identifying the 21 different geological conditions, as well as the relevant strata control aspects. In order to correctly address each of the 21 conditions, trainees should be able to specify the associated risks and control measures for each condition. Another requirement for the system was to show animations of the possible results of ignoring or not correctly addressing the geological conditions. Table 7.3 summarises the required system features.

Table 7.3: Summary of required features of the ISGC prototype.

Objectives: Identification of the most common and dangerous geological and rock abnormalities found in the Merensky and UG2 reefs and addressing them. Trainees are required to: - Correctly identify hazardous conditions. - Specify the control measures to be taken in such conditions. - Indicate the correct support for the conditions.
Target Population: All mining employees who enter areas where geological and dangerous rock conditions may be present.
Technology: - 3D simulation of underground rock texture and mining areas. - VR clips covering the required conditions and possible consequences.
Technical and media requirements: - High-quality graphics to create a virtual workplace environment. - Interactive processes that result in increased active participation and retention of information. - Multilingualism: text and audio should be available in English, Tswana, Sepedi and Xhosa. - Appropriate sound effects to increase realism of the simulation. - Training and exposure to safety hazards, yet without entering actual dangerous situations. - Randomisation of the presentation sequence of geological hazards and severity levels, in order to avoid monotonous repetitive exercises.

Administrative and Reporting Facilities:

- Trainee performance captured in a database for reference purposes.
- Report System:
    - Graphical comparisons and illustration.
    - Identification of areas where more training may be required.
    - Individual reports.
    - Group reports.
    - Periodic reports.
    - Identification of potentially high-risk employees.

7.4.2. Scene design

Even though the findings of the Usability Context Analysis, as presented in Section 6.5, informed designers and developers regarding the context in which the users work and learn, so as to support effective design and development, there is no substitute for experiencing the actual working environment and conditions. For this reason it was deemed important, prior to the production of *ISGC*, for the designers and developers who were unfamiliar with mining situations to be exposed to the actual environment. Hence an underground site visit was scheduled. This visit also allowed designers and developers to capture relevant photo and video material of prevailing geological conditions. This material was subsequently used to create realistic underground simulations. An additional challenge for the developers was to become familiar with mining terminology.

The interaction design process allows for various design ideas to be generated. Designers need to select between options, or merge multiple solutions, and re-think assumptions that may have been accepted previously. A key technique in relating design to requirements, is to make use of storyboards. These provide rough sketches of what the system will look like and are suitable for predominantly graphical interfaces, such as VR systems.

Upon completion of the underground site visit, the VR system developers in the design team designed each required scene by expressing ideas for the scene. Through brainstorming and discussions in the design team, guided by the findings of the case study regarding locations and dimensions of related aspects, final storyboards were developed by producing a series of sketches depicting the environment, characters, and a sequence of verbal dialogue and animation required for each scene. Appendix C contains an example of a storyboard used in the development of *ISGC*.

As mentioned in Section 7.3, the design of each scene in the prototype consists of two parts: a hazard identification scene and an accident scene. The hazard identification scenes are 3D scenes that show the sub-standard strata control conditions that the trainee needs to identify. Each accident scene shows possible consequences (falls of ground) that could occur for the given scenario, should the trainee not correctly support or address the condition.

7.4.3. Scene development

The *LSF* prototype was developed using the Torque game engine. In the *LSF* prototype, some of the geological conditions could be difficult to distinguish, for example, a visible crack in the rock may be a joint, a prominent joint, or a low angle joint. Therefore the *ISGC* prototype required graphics with a high level of realism in order to accurately simulate the geological conditions. This could not be achieved with the Torque game engine environment and the scenes in the prototype were subsequently developed in 3D Studio Max, which is the industry standard for animation movies.

The developers worked from the storyboards to model, texture and animate the required objects. Special visual effects were added using Adobe After Effects. Relevant underground mining sounds were added and appropriate lighting was used to simulate the effect of a miner’s cap lamp in the dark environment. Additional tools used include Adobe Photoshop, Motion Builder, Particle Illusion and Sony Vegas.

Figures 7.16 to 7.21 are extracts from the *ISGC* prototype. The scenes were developed to show geological conditions in the stope areas. *Stoping* is the process by which the orebody is broken and extracted from the working stope face for subsequent transport to the shaft and hoisting to surface. *Stope gullies* are excavation cuts in the footwall and provide the access route into stopes for people and material, and removal of ore. Broken ore is scraped down the stope face into the gully and along the gully into the box hole (a short raise in which a box, or chute, is placed to control the flow of broken rock through it). If gully conditions are permitted to deteriorate, the risk of injury is high, therefore each scene starts by first showing the gully area (see Figure 7.16). The timber support, painted mark lines and inserted roof bolts are clearly visible.

The camera then moves forward slowly down the gully until the advanced strike gully is visible, as shown in Figure 7.17. An *advanced strike gully* (ASG) is a gully that is developed ahead of the stope panel face without carrying a wide siding. This camera movement allows the trainee to inspect the gully and ASG areas for possible geological or substandard conditions.

Figure 7.16: The stope gully area.

Figure 7.17: The advanced strike gully area.

The camera then pans away into the stope face area, showing the hanging wall, face and footwall. At some stage during the scene a geological condition will be visible, which the trainee needs to identify correctly. As an example, Figure 7.18 indicates blocky ground near the face area (blocky ground is one of the geological features explained in Table 7.2).

Figure 7.18: Blocky ground in the hanging near the stope face area.

The second part of each scene portrays the possible consequences of ignoring or not correctly addressing the geological condition. As an example, Figure 7.19, which is the first part of a scene, shows a miner working in the blocky ground stope area. Figure 7.20 is a follow-up screenshot showing a rock fall that injured the miner. In a further example, shown in Figure 7.21, a miner is injured during a rock fall in an incorrectly supported area.

Figure 7.19: A miner in the unsupported area of blocky ground in the stope face area.

Figure 7.20: A fall of ground occurs, injuring the miner.

Figure 7.21: A fall of ground in an incorrectly supported area.

7.4.4. Scene evaluation

Due to users not being available to serve as participants at early stages, it was decided that the design team would use cognitive walkthrough as the approach to formative evaluation. This applied to scene evaluation, user interaction evaluation and back-end system evaluation. Cognitive walkthroughs were discussed in Section 3.6.2.2. The team went through each action, just as the user would, in order to perform a task in the system. For each action, the design was scrutinised and reviewed, noting problematic features so that alternative designs could be explored. The scenes were also shown to a group of six rock engineers at the Mine. Their feedback was incorporated in improving and fine-tuning the eventual final scenes.

7.4.5. User interaction design

Based on the user requirements, described in Section 7.4.1, the user interaction was designed according to the program flow depicted in Figure 7.22.

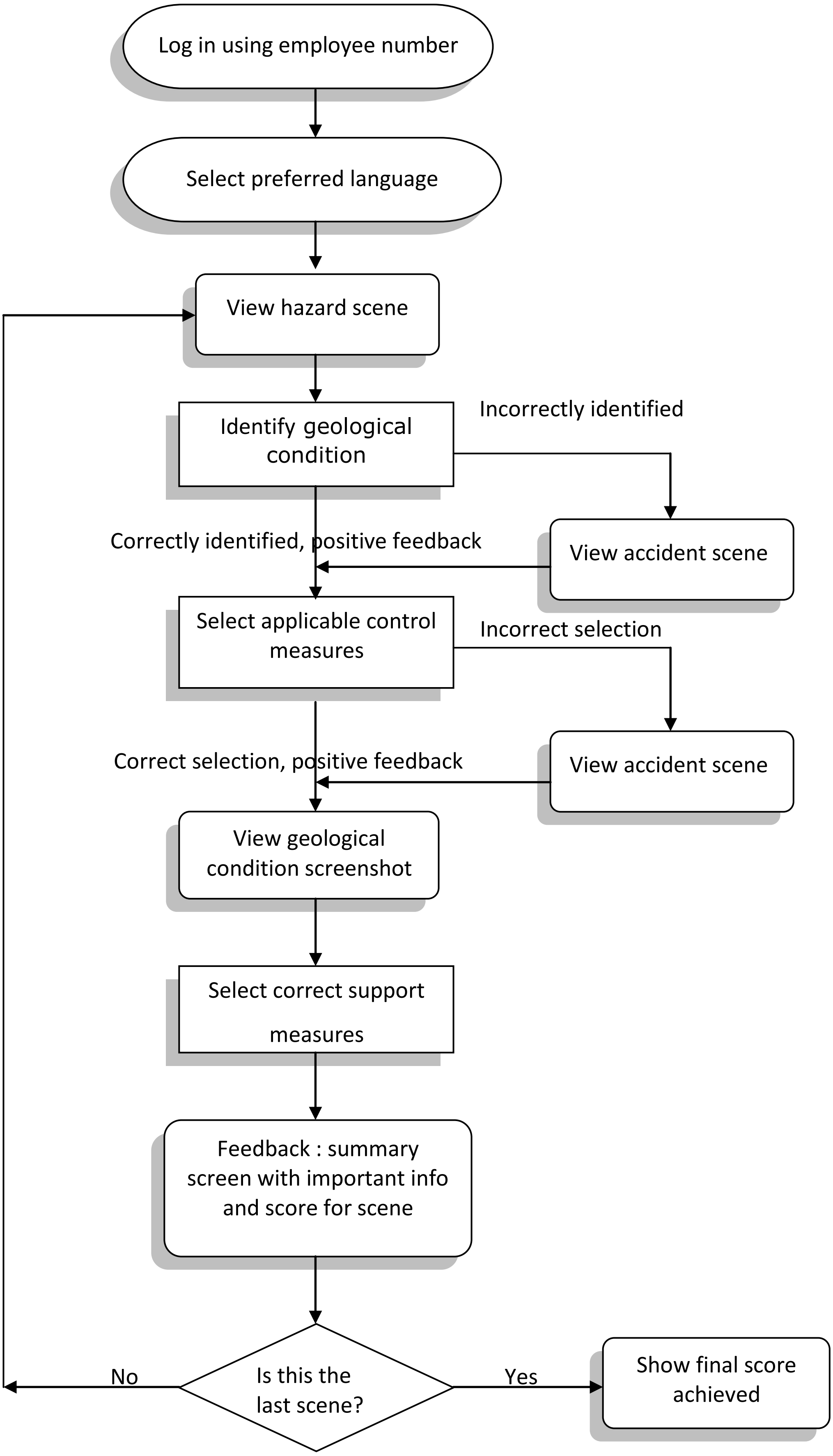


Figure 7.22: *ISGC* program flow indicating user interaction.

A detailed explanation follows, listing the steps for interacting with the system:

- The training facilitator pre-loads the employee numbers of trainees into the system database, using the administrator subsystem developed for this purpose.
- The trainee logs onto the system using his/her employee number.
- The trainee selects his/her preferred language. Available options are English, Tswana, Sepedi and Xhosa.
- An introductory video is shown to highlight the objectives of the training, as well as the functioning of the system.
- A 3D scene is shown depicting a hazardous condition. The trainee can select between three possible actions:
  - *Replay* to view the scene again,
  - *Photo*s to view photos of similar conditions, or
  - *Next* to continue to the next screen.
- Selecting *Next* will show the Hazard Identification screen with five options to choose from. The trainee needs to select the correct hazard.
- If the hazard is not identified correctly, the consequence scene is shown, which will be an appropriate fall of ground (FOG).
- If the correct hazard is selected, the trainee is now shown eight possible options and must select the correct control measures to be implemented in response to the current hazard. In some cases, more than one option could be correct, depending on the particular hazard.
- If all the correct control measures are not selected by the trainee, the FOG is shown as a possible consequence and the correct answers are shown to the trainee.
- The trainee is then shown a screenshot from the simulation and needs to select between four options regarding how this condition should be correctly supported or treated.
- The correct answers are shown, as well as the graphical representation of the correct support.
- A summary screen is shown for the hazard in question, highlighting the important information for this type of hazard and also indicating the trainee scores achieved for identification, controls and support.
- The trainee continues through the system, completing all 21 scenes and then receives a final score.
- The training facilitator uses the administrative subsystem to generate reports on trainee performance.

The learning paradigm underlying the design of the *ISGC* prototype is behaviourism (Section 3.4.1). Behaviourism aims to change trainees' behaviour by stimulus-response patterns that promote correct actions, in this case, performance that conforms to safety norms. Reinforcement should support retention of what is learned. *ISGC* supports and reinforces safe behaviour by providing positive feedback to correct performance and by showing disturbing visualisations as the possible consequences of unsafe behaviour or incorrect choices in dealing with geological hazards. This is indicated in Figure 7.22.

A database questions-and-answers spreadsheet (called a DBQA) was developed to facilitate user interaction with the system. This spreadsheet contains columns for all the text to be captured in the database for each hazard. It includes all the options shown to the trainee during hazard identification, selection of control measures and indication of correct support.

Figure 7.23 is an extract of the DBQA, showing certain columns for the brow, dome, dyke and fault hazards. The first column gives the correct hazard names. The next five columns (light blue) indicate the options that will be shown and from which the trainee should select the correct hazard during the identification question, which follows the viewing of that hazard. The next five columns show the five options which will be presented to the trainee during the selection of the control measures for a hazard. The green columns indicate the correct answers and the red columns the incorrect options. The contents of the DBQA were implemented during the back-end development phase, and a feature was built into the system to randomise the sequence of options shown to the trainee. The back-end development is discussed in Section 7.4.9.

Scene Name	Option 1	Option 2	Option 3	Option 4	Option 5	Correct Answer 1	Correct Answer 2	Wrong Answer 1	Wrong Answer 2	Wrong Answer 3	…
Brow	Dyke	Pothole	Brow	Pegmatite Vein	Pillar Robbing	Barring	Add roof bolts or mine poles	Support down dip side	Install unstressed pack	Increase mining parameters	…
Dome	Rolling reef	Blocky jointed ground	Brow	Dome	Reef in Footwall	Add roof bolts or mine poles	Install Support poles	Increase stoping width	Do not bar as it is too dangerous	Increase panel length	…
Dyke	Pegmatoid	Rolling reef	Brow	Dyke	Brittle rock	Barring	Demarcate	Increase the burden	Reduce roof bolts	Drill near vertical rounds	…
Fault	Dyke	Pothole	Brow	Scaling Rock	Fault	Leave pillar	Install roof bolts or mine poles	Support the weak side only	Increase mining parameters	Support down dip side only	…

Figure 7.23: Extract from the DBQA of the *ISGC* prototype.

7.4.6. User interaction development

Figures 7.24 to 7.29 are screenshots that depict the user interaction with the system. Figure 7.24 shows the login screen. Figure 7.25 shows the user options during the hazard identification scene. While the scene is viewed, the trainee has three options available: *Photos* to view photographs of similar conditions, a *Pause* option to pause the scene being displayed, and *Next* to continue to the next screen, on which the hazard must be identified. The trainee then needs to choose the hazard from the five options displayed in Figure 7.26.

Figure 7.24: *ISGC* login screen.

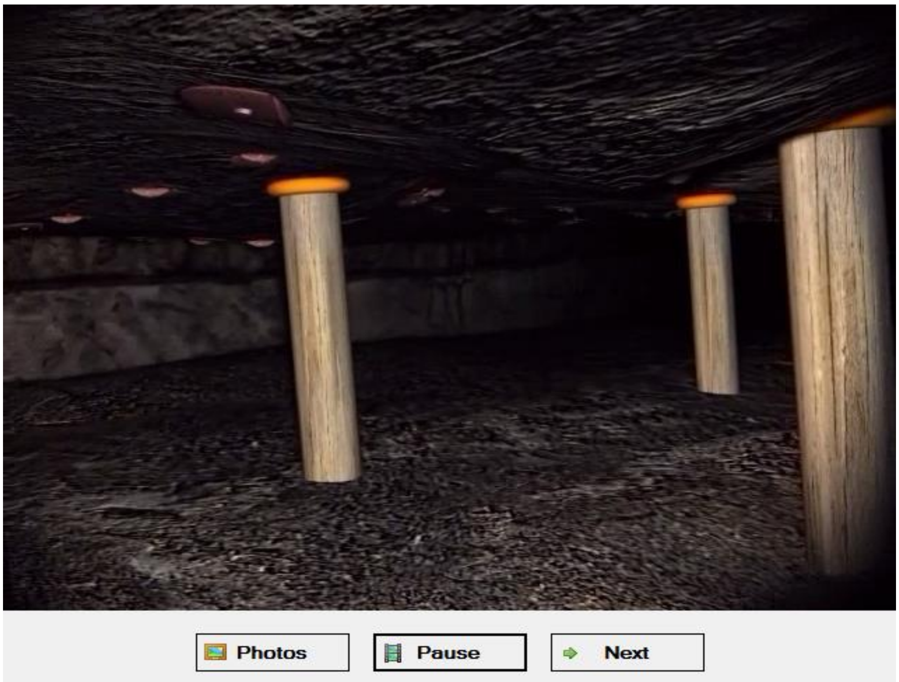


Figure 7.25: The user screen during the hazard identification phase.

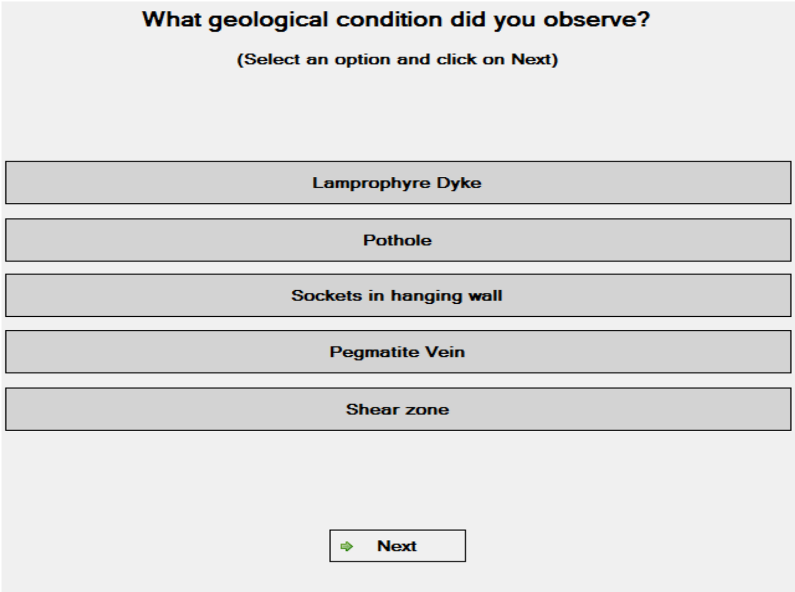


Figure 7.26: Hazard identification interaction screen, following the situation in Figure 7.25.

After the identification of the hazard, in this case a shear zone, eight possible options are given, from which the trainee selects the correct control measures to be implemented for a shear zone hazard. These options are shown in Figure 7.27. In this particular case, four answers are correct and should all be selected. Options are presented as text on grey bars, but when a trainee selects an option, the selected option background colour changes to blue. In the example in Figure 7.27, the trainee has already selected two possible answers and needs to select another two before clicking on *Next*.

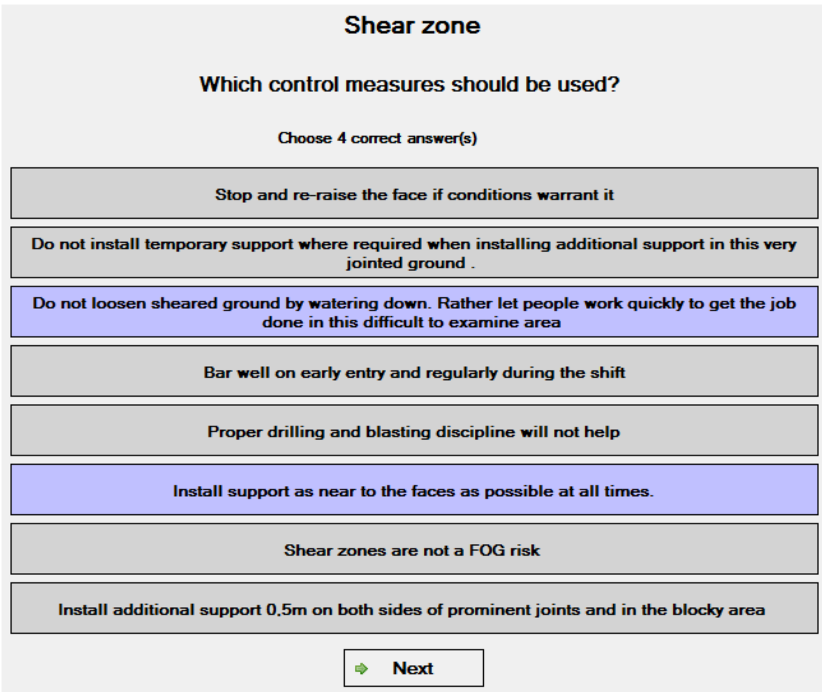


Figure 7.27: Control measure options for a shear zone.

As indicated in the program flow diagram in Figure 7.22, if the hazard is not correctly identified, or if not all the correct control measures are specified, the FOG is shown as a possible consequence. Figure 7.28 is a screenshot of the instant of the FOG, only shown to trainees who do not specify all the correct control measures. Figure 7.29 indicates how the correct answers are shown to the trainee, indicated in blue.

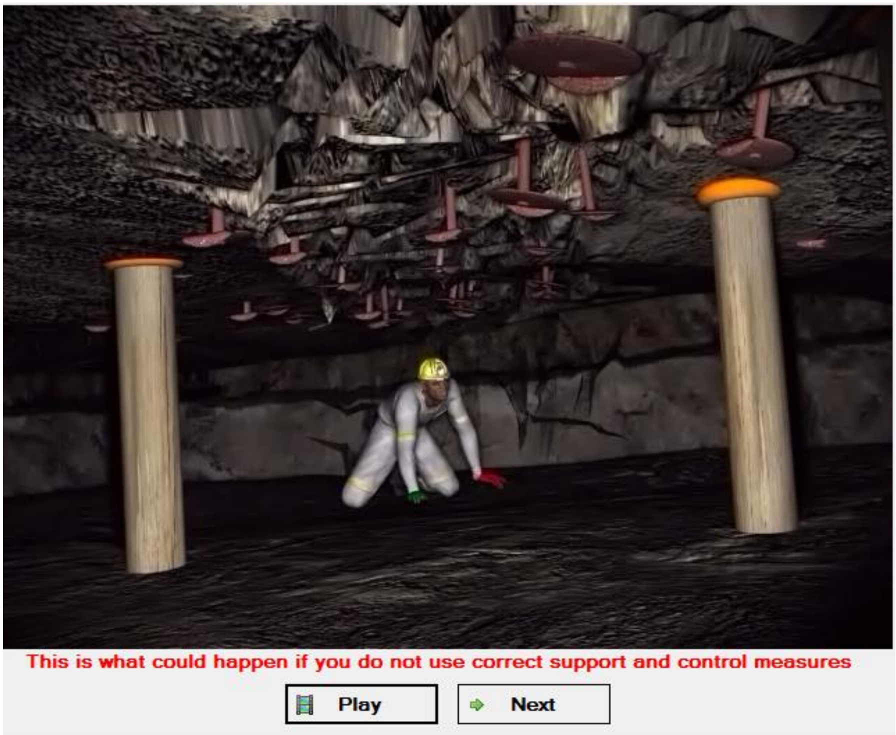


Figure 7.28: The FOG scene for a shear zone.

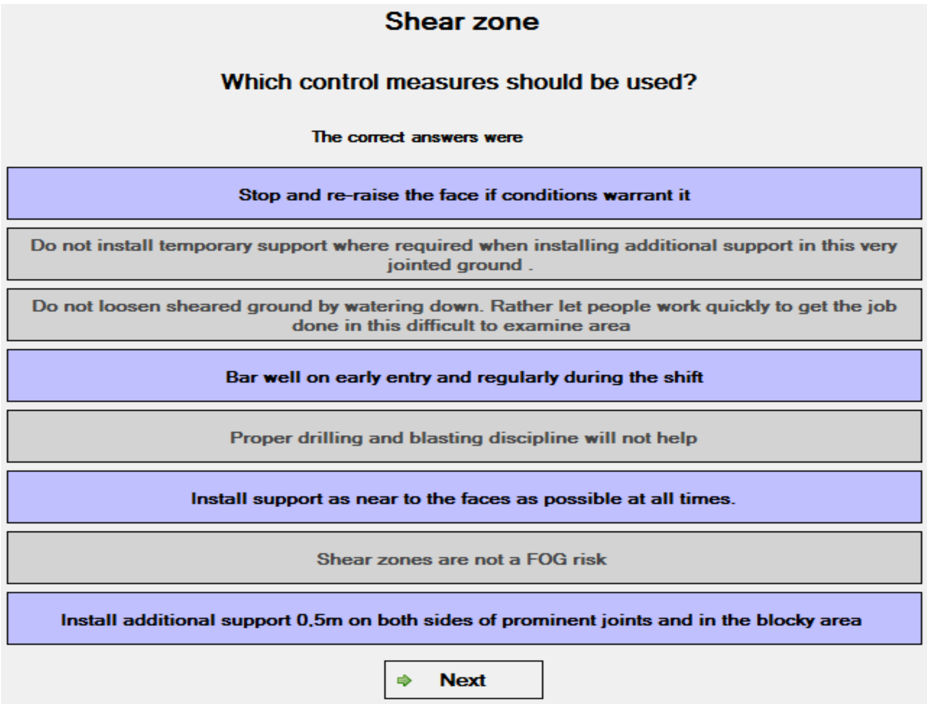


Figure 7.29: Correct answers for control measures for a shear zone.

One of the key requirements of the system is the availability of all audio and text material in different languages. This required the DBQA content, introduction video script and user feedback to be translated and recorded in Tswana, Sepedi and Xhosa. A professional translation service was used for the translations, and students familiar with these languages were used as narrators.

7.4.7. User interaction evaluation

As was the case with the screen evaluation, cognitive walkthrough was used as evaluation method. All the subject experts were involved in the design of the interaction text. It was decided to improve user feedback by adding sets of positive and negative remarks; the system would randomly select one for feedback after a trainee has submitted answers, depending on whether the answers were correct or incorrect. With this approach, the trainee would not receive the same feedback for every correct response and every incorrect response.

7.4.8. Back-end system design

Even though back-end system design is indicated in the same time phase of the extended interaction design model in Figure 7.15, it could only commence once the program flow had been finalised during interaction design. The back-end design team then decided on the appropriate programming environment in which to implement the proposed system, while the DBQA design and development was done. Thereafter, the design of the database structure, tables and indices could be done by the back-end design team.

A system such as *ISGC* can be implemented in several programming languages and databases. In the case of the *ISGC* prototype it was decided to use C# as the programming environment and MySQL as the database, due to availability of the software and programmer experience. The Indeo Video codec was used to allow full-speed video playback without requiring hardware acceleration. The back-end system design, development and evaluation were done by a team of three programmers, who were students taking the industry exposure component of their information technology qualifications.

The administration subsystem was developed in the server-side scripting language, PHP. This meant that the system executed as a browser-based system accessible via the Web.

The administration subsystem was designed to allow the training facilitator to pre-load trainee particulars in the database for authorisation purposes, so that the system could determine whether a trainee is authorised to log in.

It was envisaged that future versions of the system could make provision for a training matrix of applicable scenes according to job categories. Furthermore, there should be the capability to link the system to the mine's enterprise resource planning (ERP) software, enabling direct interaction between the *ISGC* system and the employee's training records on the ERP system.

7.4.9. Back-end system development

Apart from programming the back-end solution, the back-end team was also responsible for recording and editing the audio in the four languages, English, Tswana, Sepedi and Xhosa, as well as the capturing thereof in the correct database fields.

Several reports were developed to provide feedback on trainees' performance. Graphs were used to illustrate performance in the three categories of Identification, Control measures and Support. Individual trainee performance reports, as well as group reports per period, were developed. Facilitators could also utilise such feedback to identify potentially high-risk employees and send them for re-training before permitting them to work underground again.

7.4.10. Back-end system evaluation

A process of testing and debugging was followed to correct program errors. In the absence of completed scenes, dummy scenes were used to test the program flow.

7.4.11. System integration

Upon completion of all three simultaneous cycles of the scenes, user interaction and back-end design and development, these subsystems were integrated into one system to function as a single unit. This was done by the back-end development team.

7.4.12. System evaluation

At this stage it was important to validate that the overall system provided the functionality specified in the user requirements and that the dynamic characteristics of the system matched those required. The system was tested with real data to assess its performance.

7.4.13. System implementation

ISGC was installed at the training centre at Impala Platinum. Two facilitators were trained in using the training prototype and administration subsystem. The mine decided that all miners, shift supervisors and mine overseers returning from their annual leave should work through the prototype as refresher training before going underground again. In this case, *miner* refers to the job category and not a general mine worker. A miner has at least six years experience and is responsible for all the activities at the work face, including blasting and support. Generally, two to three miners report to the shift supervisor. It is therefore obvious that miners, shift supervisors and mine overseers should have detailed knowledge of the geological conditions prevailing at the mine.

7.4.14. User system evaluation

As determined from the context-of-use study described in Chapter Six, more than 80% of the users had not been exposed to computer-based training prior to the introduction of these prototypes. E-training is therefore a new technology in the system implementation environment and if the system is not easily usable, learners would spend excessive time trying to understand the system rather than engaging with the content. This further emphasises the need for a thorough user evaluation.

The *ISGC* system was designed from a more formal theoretical foundation than *LSF* and required a formal, theory-based evaluation. After installation of *ISGC* at the mine training centre, the system was evaluated through heuristic evaluation and a user satisfaction questionnaire. The heuristic evaluation was done by the same six experts who evaluated the *LSF* prototype, as described in Sections 5.6.6 and 5.6.7, using the *DEVREF* Evaluation Framework. The user satisfaction questionnaire was completed by 52 trainees after completion of the *ISGC* training. Results of these two major evaluations are discussed in Sections 8.4 and 8.5 respectively.

7.5. Design improvements

As indicated in Figure 7.15, the extended interaction design model proposed in this study allows for cycles of revision. In the application of the model for the design and development of the *ISGC* prototype, three cycles occurred:

1. After evaluation of each of the three design processes running simultaneously (Sections 7.4.4, 7.4.7 and 7.4.10), improvements were made to the design and development of each process.
2. After the system evaluation phase (Section 7.4.14), a number of improvements were made to the design, interaction and back-end of the system, prior to implementation of the system.
3. After user evaluation by the heuristic evaluators and the users, further improvements were required. These are detailed in Chapter Eight, which describes and discusses these evaluations.

Listed below are some of the improvements following the first two cycles described above:

- Addition of a *Risks* screen to the system. Trainees are required to select which risks are related to a particular hazard.
- Links to additional video material related to the control and support of a hazard.
- An option to view the mine’s Code of Practice (COP) document related to a hazard.
- A *Redo* option allowing the trainee to redo the training if the desired pass score was not achieved.
- A redesigned summary screen to give relevant feedback to the trainee.
- An option to view the accident scene even when the trainee had selected all the correct answers. This is accompanied by user feedback stating that this is a possible consequence which the user had avoided by correctly identifying and addressing the hazard.
- The translated text should be closely in line with the terminology used at the mine and not the pure translation.

The prototyping paradigm in software development is based on the premise that the developer builds a partially complete system in order to explore and test certain aspects of the system requirements. In the prototype discussed in this chapter, evolutionary prototyping was used, where a prototype is continually developed until it evolves into the final product.

7.6. Conclusion

This chapter presented details of a case study that analysed the circumstances relating to falls-of-ground incidents at a large platinum mine, Impala Rustenburg. This case study, which was described in Section 7.2, informed the design of a second interactive desktop VR prototype, the *Interactive Simulated Geological Conditions (ISGC)* prototype. An important conclusion resulting from the case study was that safety training relating to geological conditions should be revised and improved.

In Section 7.3 an extended interaction design lifecycle model for VR training simulations was presented. For the design of the VR training prototypes used in this study, the interaction design lifecycle model of Rogers *et al*. (2011) was extended by the researcher to make provision for three simultaneous processes, and the subsequent integration thereof into a single product.

In Section 7.4 this extended model was applied to the design of *ISGC*. Details were given on each phase of the design and development process. Formative and summative evaluations led to several design and development improvements, and subsequently, a more improved training system.

The main deliverable of this chapter is the *ISGC* prototype system. *ISGC* incorporates learning material on the identification of 21 different geological conditions, as well as the relevant strata control aspects. In order to correctly address each of the 21 conditions, trainees must specify the associated risks and control measures for each condition. *ISGC* also shows animations of the possible results of ignoring or not correctly addressing the geological conditions.

The chapter continued to address Research Subquestion 3: "What are the contextual requirements for virtual reality training systems for the mining industry?" in Section 7.2, which had already been partially addressed in Section 6.5. This chapter also addressed Research Subquestion 4: "What is an appropriate design lifecycle model for interactive desktop VR training systems?" in Section 7.3. Due to the fact that this study introduced new technology and innovative training methods into the mining industry, a suitable design lifecycle model for VR systems should include all stakeholders to ensure successful adoption. Interaction design was used as the basis for the proposed design lifecycle model due to it having the following features, which are appropriate to the context of this study:

- involvement of stakeholders,

- establishment of user requirements,
- generation of alternative designs,
- development of interactive systems, and
- evaluation of the resulting product.

The interaction design model was expanded to include three simultaneous processes for the design, development and evaluation of scenes, user interaction and the back-end system (as shown in in Figure 7.15). As stated in Section 5.5.4, a potential type of theoretical outcome of DBR is a design methodology that serves as a guideline on how to implement a design. The application of the extended interaction design lifecycle model led to the successful design, development and implementation of *ISGC* at the Mine, prompting mine management to make the improved *ISGC* compulsory training for all ex-leaves. The extended interaction design lifecycle model is therefore appropriate for designing interactive desktop VR training systems, and is a further deliverable of this chapter.

Due to e-training being a new paradigm in the mine safety training environment, a focus on the instructional design and usability of such training systems is vital to ensure effective transfer of learning content and to prevent trainees spending excessive time trying to understand the system rather than engaging with the content. This further emphasises the need for a thorough user evaluation of such systems.

The next chapter, Chapter Eight, details the heuristic evaluations and user satisfaction surveys done on both the *LSF* and *ISGC* prototypes.

Chapter Eight
Evaluation

8.1. Introduction

Using concepts from the literature and experience in the field, the researcher developed an evaluation framework for evaluating desktop VR mine safety training systems. This framework, named the Desktop Virtual Reality Evaluation Framework (*DEVREF*), was presented in Section 5.8. The *DEVREF* Evaluation Framework encompasses and emphasises aspects such as the traditional tenets of usability and factors relating to usable design of VR systems, and also addresses instructional design. Furthermore, attention is paid to context-specific issues, in this case the domain of underground mining.

The empirical research described in this chapter relates to the application of *DEVREF*, which is the primary focus of this study. To validate *DEVREF* and to use it to assess the two VR prototypes, *LSF* and *ISGC*, heuristic evaluation was applied, using the *DEVREF* criteria. As described in Section 3.6.2.3 and Section 5.7.6, heuristic evaluation is an inspection method conducted by experts and is a popular usability evaluation method for computer system interfaces, because it is quick, inexpensive, and effective at achieving broad coverage of a whole user interface. However, it may miss certain important issues that users could identify, therefore an end-user satisfaction survey was also used to evaluate the two VR systems. User-based methods are useful for establishing detailed factors such as preference, the problems that end-users encounter, and how long it takes to complete tasks. Both user satisfaction studies and heuristic evaluation methods can provide valuable insights on usability problems in the beginning, during, and at the end of the product development lifecycle.

The origin, design and development of the *LSF* and *ISGC* prototypes were discussed in Chapters Six and Seven respectively, while this chapter reports on their evaluation. Chapter Nine reports, in turn, on the meta-evaluation done on the *DEVREF* Evaluation Framework, which led to improvements to the evaluation framework itself.

The findings in this chapter are presented in detail, indicating the richness and extent of the results when evaluations are performed with *DEVREF* and the user satisfaction survey. Applying the *DEVREF* Framework not only led to improvements to the prototypes, but also resulted in the identification of inadequacies and weaknesses in the evaluation framework itself. These inadequacies and weaknesses are pointed out in discussion of findings throughout the chapter and are consolidated in Chapter Nine, which, together with the findings of the meta-evaluation, culminated in a revised and improved evaluation framework.

As indicated in Figure 8.1, the chapter is structured as follows: Sections 8.2 and 8.3 explain the design of the evaluation instruments. Section 8.4 discusses the evaluation of the *LSF* prototype as explained in Cycle 2 of the DBR process in Section 5.6. *LSF* was evaluated by heuristic evaluation using the proposed *DEVREF* Evaluation Framework, and by a user satisfaction questionnaire. Section 8.5 reports on the evaluation of the *ISGC* prototype as explained in Cycle 3 of the DBR process in Section 5.6, also by applying the *DEVREF* Framework and the user satisfaction questionnaire. Following the presentation of the *ISGC* evaluation, Section 8.6 compares the evaluation findings of the two prototypes.

It is important to point out that although both the *LSF* and *ISGC* prototypes are now in operational use as real-world outputs of the DBR process, they are still viewed as prototypes because they are refined on a regular basis.

Figure 8.2, based on Figure 5.6, indicates which parts of the research processes are covered by this chapter. The sections of DBR Cycles 2 and 3 shown in colour are relevant to this chapter, i.e. Evaluation of the prototypes (turquoise) and Reflection on the resultant findings (purple). Reflection is not covered as a separate topic in the chapter, but is incorporated in the interpretation of findings interspersed throughout the chapter. In Chapters 6 and 7 there are similar figures to Figure 8.2 reporting on other empirical work, but the figures in Chapters 8 and 9 are of particular importance. They address the core of the study in that they consider research resulting from the application of *DEVREF* (Chapter 8), and the evaluation of *DEVREF* itself (Chapter 9).

Figure 8.1 shows the layout of this chapter.

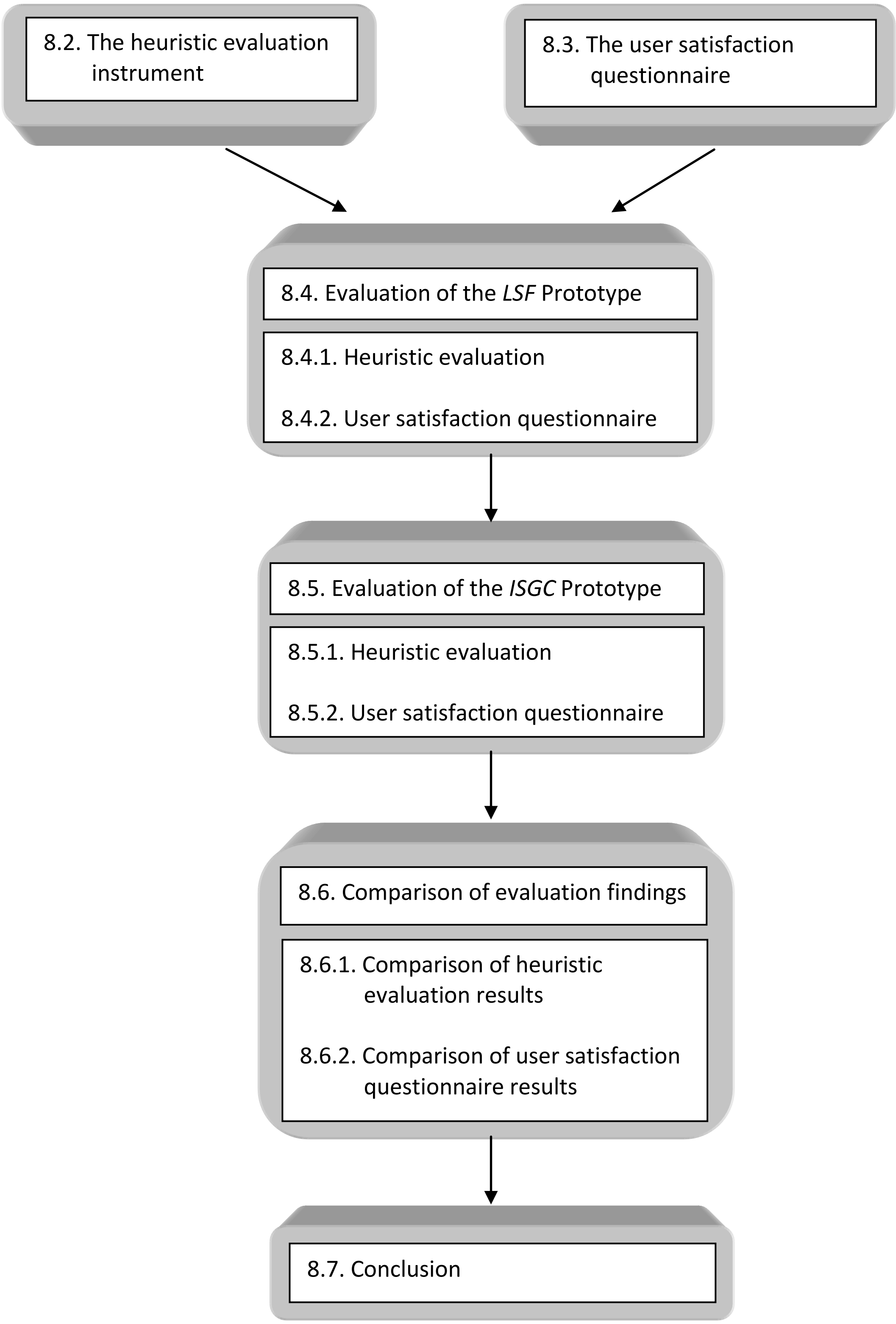


Figure 8.1: Layout of Chapter Eight.

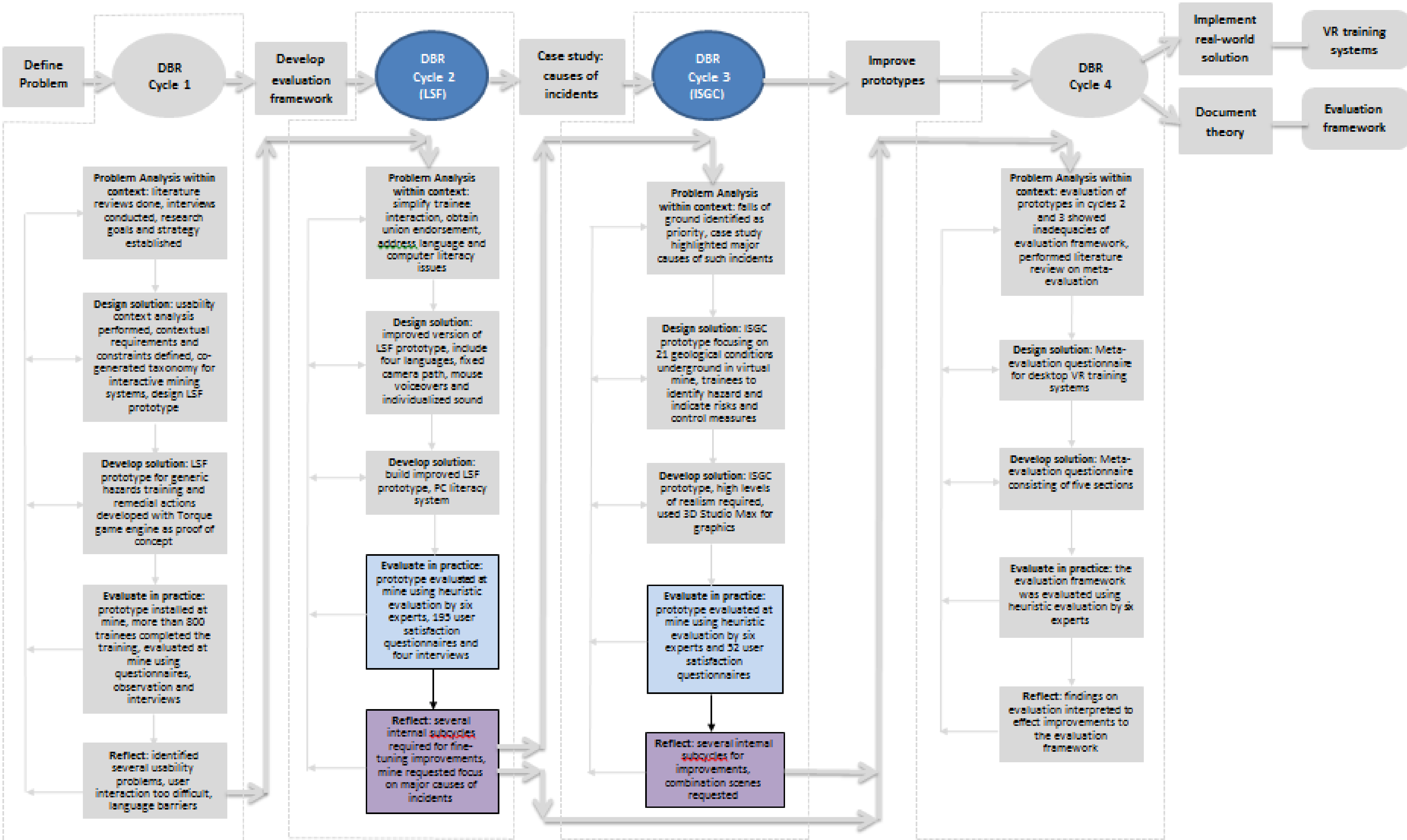


Figure 8.2: Research process flow discussed in Chapter Eight.

8.2. The heuristic evaluation instrument

Section 5.8 in the research design and methodology chapter presented *DEVREF*, a synthesised framework for heuristic evaluation of desktop VR training applications. *DEVREF* comprises four categories of heuristics, also called criteria:

- Category 1: Instructional design – includes heuristics related to pedagogical effectiveness, learning theories and multimedia learning design.
- Category 2: General usability – includes interface design and interaction, and heuristics that support the goals of usability.
- Category 3: Virtual reality system design – includes heuristics specific to the design of virtual reality systems.
- Category 4: Context-related heuristics – includes heuristics related to the content and the application domain.

For the purpose of heuristically evaluating desktop VR systems, a heuristic evaluation instrument was developed by drafting evaluation statements for each heuristic in the *DEVREF* Framework. There is a direct mapping between heuristics and evaluation statements, whereby each theoretical criterion is rephrased as a set of items for inclusion in the questionnaire. As an example, the first heuristic in the framework for the category Instructional Design is *Clear goals, objectives or outcomes.* In order to evaluate the implementation of this heuristic in a target system, the following evaluation statements were posed:

- There are clear goals, objectives or outcomes for the training program.
- Clear goals, objectives or outcomes are communicated at the beginning of the training program.
- The outcomes are measurable.

The application of each of these evaluation statements in the target system is then rated using a five-point Likert scale, ranging from *Strongly disagree* as a rating of 1 to *Strongly agree* as rating 5. Table 8.1 indicates the number of heuristics per category, as well as the number of evaluation statements for each category. In total, for the four categories, the heuristic evaluator needs to evaluate 105 evaluation statements.

At the end of each category, additional space was provided for evaluators to mention other problems that could not fit in the space provided, as well as an open-ended area for writing any further comments or elaborations. The heuristic evaluation instrument also included a consent form to be signed by the evaluator.

Table 8.1: Number of heuristics and evaluation statements per category.

Category	Number of heuristics to evaluate	Number of evaluation statements
Instructional Design	8	28
General Usability	8	28
VR System Design	8	25
Context-specific	8	24
TOTAL	**32**	**105**

A protocol for completing the heuristic evaluation questionnaire was provided to assist the heuristic evaluators, covering topics such as what was expected from them, information on the categories, important background information, and the consent form. In order to enable evaluators to judge the appropriateness of the system's usability and other aspects in an informed manner, they were required to note the following background information related to the prototypes to be evaluated.

- Application domain: Safety training for the mining industry.
- Prototypes to be evaluated: *Look, Stop and Fix (LSF)* generic hazards system and *Interactive Simulated Geological Conditions (ISGC)*.
- Target audience: *LSF* training is done by all the lower-level mine workers who work underground. Typical job positions are rock drill operator, winch operator and panel operators. The prior exposure of these trainees to computer technology ranges from very limited to none at all. The *ISGC* training is done by employees who have been promoted to higher ranks, such as shift bosses, artisans and mine captains.
- System objectives: The *LSF* prototype simulates the underground working areas, incorporating potential hazards that mine workers need to identify and to indicate possible actions that might be followed in response to each hazard. Trainees must learn to spot these potentially hazardous conditions, identify the hazards correctly, and indicate which action/s should be taken to address the situation. The *ISGC* prototype focuses on the geological conditions that may cause rock falls. Trainees have to identify the conditions correctly and specify the associated risks and control measures for each condition.
- Context of use: Both prototypes are used for refresher training of workers returning from their annual leave. Successful completion of *LSF* is compulsory before workers are allowed to work underground again. Trainees not scoring 80% after two attempts are sent for re-training.
- Program development status: Both prototypes are currently in use at several mine training centres. Annual upgrades are developed, and the results of the

evaluations described here can be used as input into iterative development, so as to improve future versions of the prototypes.

The heuristic evaluation instrument, including the protocol for completion and the evaluator consent form, is provided as Appendix B-1. This heuristic evaluation instrument was used to evaluate both the *LSF* and *ISGC* prototypes.

Section 8.4.1 discusses the findings of the *LSF* heuristic evaluation and Section 8.5.1 the findings of the *ISGC* heuristic evaluation.

8.3. The user satisfaction questionnaire

In order to evaluate user satisfaction of both the *LSF* and *ISGC* prototypes, a mixed-methods survey instrument was developed with three sections:

- Biographic details of participants.
- Questions on the use of the prototype (quantitative data).
- Open-ended questions (qualitative data).

The participants, who were authentic underground mine workers, were required to sign a consent form, stating that they willingly participated in this research, that they were aware that the findings of the evaluation would be used for research purposes and that their anonymity would be protected. The user satisfaction questionnaire, which includes the consent form, is attached as Appendix B-2.

Section 1: Biographic details

This section required each participant to state his name, age, job title, employee number and home language. Participants were requested to indicate at which mine they worked and how many years mining experience they had. Participants were also asked to indicate the highest grade they had completed at school. In order to determine their previous exposure to technology, they were asked to indicate if they had previously used each of the following: a computer, a cell phone and a bank automatic teller machine (ATM). Finally, participants had to indicate whether they had done the voluntary computer mouse pre-training prior to using the *LSF* prototype.

Section 2: Closed questions related to the prototype

This section covered various usability aspects, including ease of use, acceptability, performance, learnability, efficiency, authenticity, memorability and method of choice. A five-point Likert scale was used to indicate participant responses ranging from the extreme positive to the extreme negative, for instance, *Very much*, *Much, Average*, *Not much*, and *Not at all*.

Section 3: Open-ended questions

Participants were given an opportunity to comment on the features of the program, aspects that should be improved, possible problems encountered, and other possible training that could also be done using desktop VR. Participants could write their own answers in the space provided.

Since different users were involved in the two prototypes, the same user satisfaction survey instrument was used to evaluate both the *LSF* and *ISGC* prototypes. Section 8.4.2 discusses the findings of the *LSF* evaluation and Section 8.5.2 the findings of the *ISGC* evaluation.

8.4. Evaluation of the Look, Stop and Fix prototype

The *Look, Stop and Fix* (*LSF*) prototype was evaluated using the heuristic evaluation instrument described in Section 8.2 and the user satisfaction questionnaire described in Section 8.3. These evaluations form part of the research activities of Cycle 2 of the four cycles of the synthesised design-based research model, as described in the Research Design section in Section 5.6 (and also indicated in Figure 8.2).

The *LSF* prototype is described in detail in Section 6.6, along with illustrative screen shots. This prototype training system focuses on hazard recognition and training in remedial actions for conventional mining. The prototype simulates the underground working areas, incorporating potential hazards that mine workers need to identify, as well as indicating the appropriate actions to be followed in response to each hazard. The target group of *LSF* is all the lower-level mine workers who work underground.

Having heuristic evaluators with expertise in all the relevant areas would, of course, be ideal, but such evaluators would generally be hard to find. In this research, it was decided to approach evaluators who preferably had expertise in more than one evaluation category. The heuristic evaluation was done by six experts: two usability experts, two mining training experts and two VR development experts. The two usability experts (Evaluators A and B) both have completed master's degrees on usability and are currently busy with PhD studies. Both of them have previous experience in usability evaluation and instructional design, but have had very limited to no exposure to VR system design and the mine safety training environment. The mining training experts (Evaluators C and D) are experts in instructional design in the mining training environment and both are heads of their respective departments. Evaluator C also has some knowledge of VR system design. The VR developers (Evaluators E and F) have been involved in such development for the past seven years. Evaluator E has experience in usability, VR system design and mine training system development, with Evaluator F being an expert in VR system design for the mining environment. Table 8.2 indicates the overlapping expertise of the heuristic evaluators used.

Table 8.2: Expertise of the heuristic evaluators in the categories of the DEVREF Framework.

Framework category	Usability experts		Mining training experts		VR development experts	
	Evaluator A	Evaluator B	Evaluator C	Evaluator D	Evaluator E	Evaluator F
Instructional Design	X	X	X	X		
System Usability	X	X			X	
VR system design			X		X	X
Mining expertise			X	X	X	X

For the purpose of the evaluations, the *LSF* prototype was installed on the researcher's laptop. The researcher visited each heuristic evaluator separately for an evaluation session, and the evaluators had the opportunity to interact with the prototype prior to conducting the evaluation. They also had continuous access to the prototype during the evaluation in order to support rating the evaluation statements. The two VR experts evaluated the prototype independently using the template in Appendix B-1.

The second evaluation instrument, the user satisfaction questionnaire (Appendix B-2), was completed by 195 trainees after completion of the *LSF* training. This was done over a period of three weeks at the training centre of a large platinum mine near Rustenburg, North West Province. The researcher was assisted by a Tswana-speaking training facilitator regarding the logistical aspects of the trainees completing the questionnaire.

Section 8.4.1 describes the findings of the heuristic evaluation of the *LSF* prototype, and Section 8.4.2 details the findings of the user satisfaction questionnaire.

8.4.1. Findings of the heuristic evaluation of *LSF*

The findings of the heuristic evaluation (HE) are presented in tabular format, detailing the criteria evaluated and each evaluator's response in terms of the rating (1 to 5). The evaluators rated each heuristic on a five-point scale, with 1 representing *strongly disagree* and 5 *strongly agree*. Tables 8.3 to 8.6 provide summaries of the evaluation results depicting the individual responses of all six evaluators for each evaluation statement (e.g. 1.1, 1.2, 1.3), an additional column containing the average rating of the six evaluators for that statement, additional rows containing the average per evaluator for each criterion, and the category average per evaluator. The average rating for each criterion is indicated next to the evaluator averages for each category, presented on a light grey background. For clarity, different colours are used to indicate the responses of the evaluators, based on their backgrounds. The two usability experts' responses are indicated in red (evaluators A and B), the mining experts' in green (evaluators C and D) and the VR experts' in blue (evaluators E and F).

The four tables respectively present the results for the four different categories in the framework:

- Table 8.3 summarises the results of Category 1 – Instructional Design heuristics,
- Table 8.4 covers Category 2 – General Usability heuristics,
- Table 8.5 Category 3 – VR System Design heuristics, and
- Table 8.6 Category 4 – Context-specific heuristics.

Each table is followed by discussion of the tabulated data.

8.4.1.1. *LSF* heuristic evaluation results for Category 1

Table 8.3: Heuristic evaluation results of LSF for Category 1: Instructional Design Heuristics.

CRITERIA	Evaluator rating						
	Usability experts		Mining experts		VR experts		
	A	B	C	D	E	F	Ave
1. Clear goals, objectives or outcomes							
1.1. There are clear goals, objectives or outcomes for the training program.	4	4	4	4	4	4	4

1.2. Clear goals, objectives or outcomes are communicated at the beginning of the training program.	5	2	4	4	4	4	3.8
1.3. The outcomes are measurable.	5	5	5	4	2	3	4
Criterion average per evaluator	4.7	3.7	4.3	4	3.3	3.7	3.9
2. Instructional assessment							
2.1. The program provides assessment opportunities that are aligned with the objectives or outcomes.	4	5	5	4	3	4	4.2
2.2. The assessment opportunities will serve to enhance trainees' performance.	5	5	5	3	4	4	4.3
Criterion average per evaluator	4.5	5	5	3.5	3.5	4	4.3
3. Feedback to user responses							
3.1. The training program provides trainees with constructive and supportive feedback on their performance.	5	4	5	4	3	4	4.2
3.2. The feedback is relevant to the training content.	5	4	5	5	4	4	4.5
3.3. The feedback informs the trainee regarding his level of achievement in the training program.	5	5	5	5	3	3	4.3
3.4. The feedback indicates incorrect responses and provides information on the correct responses.	5	5	5	5	3	3	4.3
Criterion average per evaluator	5	4.5	5	4.8	3.3	3.5	4.3
4. Motivation and creativity							
4.1. The system supports intrinsic motivation by providing challenges to trainees and encouragement when errors are made.	5	4	5	4	3	4	4.2
4.2. The program captures the trainee's attention early and retains it throughout.	5	5	4	5	4	4	4.5
4.3. This training program increases trainees' confidence by providing them with reasonable opportunities to accomplish the objectives successfully.	4	5	5	4	4	4	4.3
4.4. The program engages trainees by its relevant content.	5	4	5	4	4	5	4.5
4.5. The program engages trainees by its interactivity.	4	4	5	5	4	4	4.3
Criterion average per evaluator	4.6	4.4	4.8	4.4	3.8	4.2	4.4
5. Differences between individual users							
5.1. The system takes account of linguistic and cultural differences by allowing trainees to select between different languages.	5	4	5	5	5	5	4.8
5.2. The system caters for trainees with different levels of expertise regarding the content.	4	4	5	2	2	2	3.2
5.3. The system caters for trainees with different levels of computer experience.	4	3	5	4	2	3	3.5
Criterion average per evaluator	4.3	3.7	5	3.7	3	3.3	3.8
6. Reduction of extraneous processing in working memory							
6.1. The training program effectively uses signalling to highlight essential issues, such as restating important points, using headings for important points, or stressing them in audio.	5	5	5	4	2	4	4.2

6.2. Redundancy is avoided by not presenting unnecessary information.	5	2	5	4	4	5	**4.2**
6.3. Redundancy and overload are avoided by not reiterating the same material in multiple modes (.e.g. the program presents information using pictures and spoken words, rather than presenting it in pictures, spoken words, and printed words).	4	2	5	5	3	4	**3.8**
Criterion average per evaluator	**4.7**	**3**	**5**	**4.3**	**3**	**4.3**	**4.1**
7. Fostering of germane cognitive load							
7.1. The training program supports the formation of mental schema by explaining where newly acquired knowledge fits into the bigger picture.	4	4	5	4	3	3	**3.8**
7.2. The system encourages encoding of the training content into long-term memory by presenting questions after each learning segment.	5	4	4	5	4	4	**4.3**
7.3. Sufficient scaffolding support is provided (in the form of hints, prompts and feedback) to help trainees achieve training goals.	5	4	4	4	2	3	**3.7**
7.4. The training program presents narration in a colloquial conversational style.	5	4	4	5	3	3	**4**
7.5. The training program prompts trainees to link concrete example information for each problem category to more abstract information.	5	4	5	4	2	2	**3.7**
Criterion average per evaluator	**4.8**	**4**	**4.4**	**4.4**	**2.8**	**3**	**3.9**
8. Appropriate intrinsic cognitive load							
8.1. Working through the training program does not cause trainees to split their attention between multiple sources of visual information.	5	4	4	5	5	4	**4.5**
8.2. The program enhances retention by presenting information in learner-paced segments, rather than as a continuous presentation.	5	4	4	5	5	5	**4.7**
8.3. The system effectively supports dual channel processing of simultaneous visual and verbal material.	5	4	5	5	4	4	**4.5**
Criterion average per evaluator	**5**	**4**	**4.3**	**5**	**4.7**	**4.3**	**4.6**
Category average per evaluator	**4.7**	**4**	**4.7**	**4.3**	**3.4**	**3.8**	
Overall average rating for Category 1							**4.2**

An obvious observation from the information presented in Table 8.3 is that there is no regular consensus pattern between the evaluators. The only evaluation statement where all six evaluators allocated the same rating is 1.1. For each of the other evaluation statements there was some difference of opinion, albeit mostly minor differences. The differences tend to indicate the evaluators' varying perspectives.

In many cases the two VR experts (indicated in blue) assigned lower ratings than the other evaluators. This is also clear from the second last row in the table 'Category average per evaluator', where the averages of the two VR experts are 3.4 and 3.8 respectively. It holds particularly true for Criterion 7 (Fostering of germane cognitive load). Regardless of this, the average rating for the instructional design category is 4.2, which is quite close to the maximum score of 5 (*Strongly agree*). When the VR experts were asked why they assigned low ratings, they both indicated their belief that VR technology could have been applied better to address the issues covered by this criterion than the way it was done in *LSF*.

The highest average score (4.8) was allocated to Evaluation Statement 5.1: 'The system takes account of linguistic and cultural differences by allowing trainees to select between different languages'. This is due to *LSF* making provision for four languages: English, Tswana, Xhosa and Sepedi. A trainee selects a language of choice at the start of the program and then all further interaction, text and narration, is presented in that language.

Even though space was provided after each category for additional open-ended comments, few comments were received. One of the mining expert evaluators wrote a remark linked to Evaluation Statement 5.1, cautioning that, although the prototype allows trainees to select between different languages, it is important not to translate any terminology related to risk management that could cause confusion in another language if such terminology was previously mastered in English. The same evaluator also pointed out that there are many other hazards at the workplace not included in the *LSF* prototype system and that care should be taken to avoid a possible perception that these were the only hazards to look out for.

Two other comments received were that the word 'performance' in 2.2 could also refer to 'knowledge', and that the name of the system could be included on the welcome screen.

8.4.1.2. *LSF* heuristic evaluation results for Category 2

Table 8.4: Heuristic evaluation results of LSF for Category 2: General Usability Heuristics.

CRITERIA	Evaluator rating						
	Usability experts		Mining experts		VR experts		
	A	B	C	D	E	F	Ave
1. Functionality							
1.1. The interface provides the level of functionality the user requires to complete a task.	5	4	5	5	4	4	4.5
1.2. The interface provides adequate back button functionality to return to a previous screen.	4	5	5	5	3	2	4

1.3. Icons, labels and symbols are intuitive and meaningful to trainees, bearing in mind the level of trainee context and experience.	5	2	5	5	2	3	**3.7**
Criterion average per evaluator	**4.7**	**3.7**	**5**	**5**	**3**	**3**	**4.1**
2. User guidance							
2.1. The interface provides clear indications of what the next required action will be.	4	5	5	4	3	4	**4.2**
2.2. Help for operating the program is accessible at any time and appropriate.	1	2	2	2	1	1	**1.5**
2.3. Trainees receive clear instructions on how to use the training program.	5	3	4	5	3	3	**3.8**
2.4. Guidance to solve problems is given in the form of examples, diagrams, videos or photographs.	4	2	5	2	3	3	**3.2**
Criterion average per evaluator	**3.5**	**3**	**4**	**3.3**	**2.5**	**2.8**	**3.2**
3. Consistency							
3.1. There is consistency in the sequence of actions taken in similar situations.	5	5	5	5	5	5	**5**
3.2. There is consistency in the use of images, prompts, screens, menus, colours, fonts and layouts.	5	5	5	5	5	5	**5**
3.3. Objects, options, and permissible actions are visible so that users do not have to remember instructions.	5	5	5	4	4	5	**4.7**
3.4. Different screens that have similar operations, use similar elements for achieving similar tasks.	5	5	5	4	3	4	**4.3**
Criterion average per evaluator	**5**	**5**	**5**	**4.5**	**4.3**	**4.8**	**4.8**
4. Error Correction							
4.1. Error messages are expressed in plain language.	4	4	5	4	3	4	**4**
4.2. Learners are provided with the necessary help to recover from cognitive errors.	5	4	5	4	2	2	**3.7**
4.3. Error messages indicate precisely what the problem is and give simple, constructive, specific instructions for recovery.	5	4	4	3	2	2	**3.3**
Criterion average per evaluator	**4.7**	**4**	**4.7**	**3.7**	**2.3**	**2.7**	**3.7**
5. System Status							
5.1. The training program keeps the trainee informed about what is going on through constructive, appropriate and timely feedback.	5	4	5	4	3	3	**4**
5.2. For every action taken by the trainee, there is a visual or audio response by the training program so that learners can see and understand the results of their actions.	5	4	5	5	3	4	**4.3**
5.3. The program responds to actions initiated by the user and there are no surprise actions from the system's side.	5	4	5	5	4	4	**4.5**
Criterion average per evaluator	**5**	**4**	**5**	**4.7**	**3.3**	**3.7**	**4.3**
6. Aesthetics							
6.1. The screens are pleasing to look at.	5	2	4	4	2	3	**3.3**

6.2. The buttons and selections are of an adequately viewable size.	5	4	5	4	3	4	**4.2**
6.3. The text is of an adequately viewable size.	5	4	5	4	4	3	**4.2**
6.4. There is not too much content or information on the screens.	5	5	5	5	5	4	**4.8**
Criterion average per evaluator	5	3.8	4.8	4.3	3.5	3.5	**4.1**
7. Error Prevention							
7.1. The training program is designed in such a way that the learner cannot easily make serious errors.	5	5	5	3	4	4	**4.3**
7.2. When the learner makes an error, the system responds with an error message.	5	5	5	4	4	5	**4.7**
7.3. Trainees can recognise situations where errors are due to the way they provided input, and not due to incorrect content in their response.	5	5	5	4	3	3	**4.2**
7.4. The system is robust and reliable throughout.	5	2	5	3	4	5	**4**
Criterion average per evaluator	5	4.3	5	3.5	3.8	4.3	**4.3**
8. Interactivity							
8.1. The training program uses clear and simple terminologies that support trainees in understanding how to interact with the system.	4	5	5	4	4	4	**4.3**
8.2. The program provides interactions that support trainees in learning the necessary content.	5	5	5	4	3	4	**4.3**
8.3. Working through the program requires regular trainee interactivity to maintain attention and facilitate comprehension.	5	5	5	4	5	5	**4.8**
Criterion average per evaluator	4.7	5	5	4	4	4.3	**4.5**
Category average per evaluator	4.7	4.1	4.8	4.1	3.4	3.6	
Overall average rating for Category 2							**4.1**

In Table 8.4, it is once again the case that the average ratings of the two VR experts, namely 3.4 and 3.6 respectively, are lower than those of the other four evaluators. The category average is 4.1, even though two of those evaluators (A and C) had very high personal averages, namely 4.7 and 4.8 respectively, as indicated in the penultimate row 'Category average per evaluator'. In particular, the evaluation statements of Criterion 4 (Error correction) were ranked lower by the two VR experts.

In terms of Criterion 6 (Aesthetics), it is noteworthy that Evaluation Statement 6.1 'The screens are pleasing to look at', drew ratings varying from 2 to 5 from the different evaluators, demonstrating their varying perspectives – what was pleasing to some was not necessarily pleasing to others.

Evaluation Statement 7.4, 'The system is robust and reliable throughout', also drew a range of responses varying from 2 to 5.

Two comments in the spaces for open-ended responses indicated that a help button should be available to the trainees throughout the program flow. This is confirmed by the very low rating of Evaluation Statement 2.2, with an average of only 1.5. In fact, the low average rating of 3.2 for Criterion 2 (User guidance) as a whole, indicates clear shortcomings in *LSF* in this area. It was also mentioned that a touch screen type interface might be more user-friendly than mouse interaction.

8.4.1.3. *LSF* heuristic evaluation results for Category 3

Table 8.5: Heuristic evaluation results of LSF for Category 3: VR System Design Heuristics.

CRITERIA	Evaluator rating						
	Usability experts		Mining experts		VR experts		
	A	B	C	D	E	F	Ave
1. User control							
1.1. The user is able to interact with, or control, the virtual environment in a natural manner.	5	3	5	3	2	2	3.3
1.2. Responses from the environment to the participant's control actions and movements, are perceived as immediate or close-to-immediate.	5	4	5	4	4	4	4.3
1.3. The system permits easy reversal of actions.	4	2	3	1	1	2	2.2
1.4. Trainees are able to exit the system at any time when they need to do so.	2	1	3	3	1	1	1.8
Criterion average per evaluator	4	2.5	4	2.8	2	2.3	2.9
2. Multimodal System output / feedback							
2.1. The effect of the trainee's actions on objects in the virtual environment, is immediately visible and conforms to the laws of physics and the trainee's perceptual expectations.	4	4	5	4	2	3	3.7
2.2. The visual representation of the virtual world maps to the trainee's normal perception of that environment.	5	4	5	4	3	3	4
2.3. Distortions are not noticeable in visual images.	4	1	4	1	3	3	2.7
2.4. Audio is integrated seamlessly into user task activity.	5	5	5	4	3	4	4.3
2.5. Audio information is meaningful and timely.	5	5	5	4	4	4	4.5
Criterion average per evaluator	4.6	3.8	4.8	3.4	3	3.4	3.8
3. Presence							
3.1. Users feel as if they are part of the virtual environment and not isolated from it.	4	4	4	4	3	4	3.8
3.2. The virtual environment experience is consistent with similar real-world experiences.	5	4	5	3	3	4	4
Criterion average per evaluator	4.5	4	4.5	3.5	3	4	3.9

4. Orientation							
4.1. Users do not find it difficult to maintain knowledge (or 'awareness') of their location while moving through the virtual environment.	4	4	5	4	4	4	**4.2**
4.2. The virtual environment includes appropriate spatial labels and landmarks to assist user orientation.	4	4	5	4	4	5	**4.3**
4.3. It is clear to the user how to exit the virtual environment.	2	1	3	3	1	1	**1.8**
Criterion average per evaluator	**3.3**	**3**	**4.3**	**3.7**	**3**	**3.3**	**3.4**
5. Navigation							
5.1. Is it easy for users to move and reposition themselves in the virtual environment.	3	1	3	4	1	2	**2.3**
5.2. Ways of navigating are consistent throughout the system.	5	5	5	4	4	4	**4.5**
Criterion average per evaluator	**4**	**3**	**4**	**4**	**2.5**	**3**	**3.4**
6. Object interaction: selection and manipulation							
6.1. Input devices are easy to use and easy to control.	5	5	5	4	2	3	**4**
6.2. Object interactions are designed realistically to reproduce real-world interaction.	5	5	5	4	3	3	**4.2**
6.3. The system provides the ability to rotate 3D objects and increase detail levels when necessary for task performance.	2	1	1	1	1	1	**1.2**
Criterion average per evaluator	**4**	**3.7**	**3.7**	**3**	**2**	**2.3**	**3.1**
7. Fidelity							
7.1. The simulations in the system are accurate.	4	2	4	3	2	3	**3**
7.2. The objects in the virtual environment move in a natural manner.	4	2	4	1	1	3	**2.5**
7.3. The virtual environment displays adequate levels of realism.	5	3	4	2	3	3	**3.3**
7.4. High-fidelity graphics are used where required.	3	1	3	1	1	2	**1.8**
Criterion average per evaluator	**4**	**2**	**3.8**	**1.8**	**1.8**	**2.8**	**2.7**
8. Various user modes							
8.1. The system employs various user modes to cater for a range of users from novices to experts.	2	1	4	1	1	2	**1.8**
8.2. The system provides various user-guidance modes, e.g. Free mode, Presentation mode, Guided mode and Discovery mode.	3	1	1	1	1	2	**1.5**
Criterion average per evaluator	**2.5**	**1**	**2.5**	**1**	**1**	**2**	**1.7**
Category average per evaluator	**4**	**2.9**	**4**	**2.9**	**2.3**	**2.9**	
Overall average rating for Category 3							**3.2**

Table 8.5 presents the results of the heuristic evaluation of the VR system design heuristics. Once again, the lower average ratings of the two VR experts, 2.3 and 2.9 respectively, indicate that several of the design aspects in *LSF* could be improved. This is also confirmed by the fact that this category's average (3.2) is considerably lower than

the other three categories evaluated (4.1, 4.2 and 3.9 for Categories 1, 2 and 4 respectively). Evaluators A and C rated the evaluation statements an average of 4, but this can be attributed to these evaluators having very little VR experience.

Specifically, the following evaluation statements were rated less than 3 and the issues should be addressed:

- 1.3: The system permits easy reversal of actions (rating of 2.2).
- 1.4: Trainees are able to exit the system at any time when they need to do so (1.8).
- 2.3: Distortions are not noticeable in visual images (2.7).
- 4.3: It is clear to the user how to exit the virtual environment (1.8).
- 5.1: It is easy for users to move and reposition themselves in the virtual environment (2.3).
- 6.3: The system provides the ability to rotate 3D objects and increase detail levels when necessary for task performance (1.2).
- 7.2: The objects in the virtual environment move in a natural manner (2.5).
- 7.4: High-fidelity graphics are used where required (1.8).
- 8.1: The system employs various user modes to cater for a range of users from novices to experts (1.8).
- 8.2: The system provides various user-guidance modes, e.g. Free mode, Presentation mode, Guided mode and Discovery mode (1.5).

In summary, the quality of graphics and overall realism should be improved. Action reversals and exiting the system are problematic and *LSF* does not make sufficient provision for user modes and user-guidance modes. Only one additional comment was received for this category: 'Graphics are primitive and not very true to life'. This also confirms the poor ratings for Evaluation Statements 2.3 and 7.4.

8.4.1.4. *LSF* heuristic evaluation results for Category 4

Table 8.6: Heuristic evaluation results of LSF for Category 4: Context-specific Heuristics.

CRITERIA	Evaluator rating						
	Usability experts		Mining experts		VR experts		
	A	B	C	D	E	F	Ave
1. Authentic tasks							
1.1. The training system supports particular work practices in the context of their natural work environment.	5	4	5	4	4	5	4.5

1.2. The system is customised according to learner-specific needs and the relevance of the curriculum.	4	4	5	5	4	5	4.5
1.3. The program includes tasks applicable to the actual job context of the trainee.	5	4	3	4	4	4	4
Criterion average per evaluator	4.7	4	4.3	4.3	4	4.7	4.3
2. Appropriate reference materials							
2.1. The system includes additional reference materials, providing information to trainees on standard operating procedures used in the application domain.	2	1	1	1	2	3	1.7
2.2. The reference materials included in the system are relevant to the problem scenarios.	2	1	1	1	2	3	1.7
2.3. The reference materials are at a level appropriate to the trainees.	2	1	1	1	2	4	1.8
Criterion average per evaluator	2	1	1	1	2	3.3	1.7
3. Comprehensive scope							
3.1. The learning material in the program covers all the vital aspects relating to the topics being addressed.	4	4	2	5	2	4	3.5
3.2. The training also covers possible consequences of trainees not applying the learning material correctly in their work place.	5	4	4	5	4	5	4.5
Criterion average per evaluator	4.5	4	3	5	3	4.5	4
4. Adaptive design							
4.1. The design of the training system is adaptive to changes in site practices.	5	4	4	4	4	4	4.2
4.2. The system refers to the latest current standard operating procedures.	4	4	5	4	3	5	4.2
4.3. The system randomises assessment details such as questions and multiple choice answers when presenting assessment opportunities to trainees.	2	5	1	1	3	3	2.5
Criterion average per evaluator	3.7	4.3	3.3	3	3.3	4	3.6
5. Appropriate record keeping							
5.1. The system maintains student records and assessment results.	5	5	5	5	5	5	5
5.2. The system monitors and displays student progress.	5	5	5	5	4	4	4.7
5.3. The system ensures legal compliance in the application domain by capturing detailed individual performance data.	4	4	5	4	4	5	4.3
Criterion average per evaluator	4.7	4.7	5	4.7	4.3	4.7	4.7
6. Trainee preparedness							
6.1. Trainees are shown how to use the software prior to doing the training program.	5	3	5	5	4	4	4.3
6.2. PC literacy pre-training is available to trainees not comfortable with using computers for training.	5	3	5	5	4	5	4.5
Criterion average per evaluator	5	3	5	5	4	4.5	4.4

7. Relevant subject matter							
7.1. The subject matter matches the goals and objectives of the training program.	**5**	**4**	**5**	**4**	**4**	**5**	**4.5**
7.2. The subject matter is presented in an appropriate content structure.	**5**	**4**	**5**	**4**	**4**	**4**	**4.3**
7.3. The information provided in the program is accurate.	**5**	**4**	**5**	**4**	**4**	**5**	**4.5**
7.4. The system 'speaks the trainee's language' by using terms, phrases, symbols and concepts familiar to the trainee and common to the application domain.	**4**	**4**	**3**	**4**	**4**	**5**	**4**
7.5. The level of language use, in terms of grammar and style, is applicable to the target audience.	**4**	**4**	**4**	**5**	**4**	**5**	**4.3**
Criterion average per evaluator	**4.6**	**4**	**4.4**	**4.2**	**4**	**4.8**	**4.3**
8. Understandable and meaningful symbolic representation							
8.1. Symbols, icons and terminology used to represent concepts and objects are used consistently throughout the program.	**5**	**4**	**5**	**5**	**4**	**4**	**4.5**
8.2. Symbols, icons and terminology used are intuitive within the context of the task.	**5**	**4**	**5**	**4**	**3**	**3**	**4**
8.3. Metaphors used correspond to real world objects or concepts.	**5**	**4**	**3**	**4**	**2**	**3**	**3.5**
Criterion average per evaluator	**5**	**4**	**4.3**	**4.3**	**3**	**3.3**	**4**
Category Average per evaluator	**4.3**	**3.7**	**3.8**	**3.9**	**3.5**	**4.3**	
Overall average rating for Category 4							**3.9**

Even though the results of the *LSF* evaluation of the context-specific heuristics yielded a 3.9 average (as indicated in Table 8.6), the prototype performed well with regard to most criteria. The lower average is mainly due to poor ratings for Criterion 2 (Appropriate reference materials), which resulted in a very low average of 1.7 for this criterion. Specifically, in the mining industry most mines have Standard Operating Procedures (SOP's) as well as procedures on risks and hazards, called Fatal Risk Control Protocols (FRCPs) or Major Hazard Management Plans (MHMPs). To improve access to reference resources, *LSF* should include a link enabling trainees to access relevant reference materials, providing information to trainees on standard operating procedures in the application domain and the appropriate risk procedures.

The only other evaluation statement that yielded a low average (2.5) was 4.3: 'The system randomises assessment details such as questions and multiple-choice answers when presenting assessment opportunities to trainees'. Here the ratings varied from 1 to 5, that is, from no randomisation to full randomisation. In fact, the evaluators did not know whether randomisation had been implemented or not. The only way for them to determine this, would have been to work through the complete system several times to specifically check the order of questions and answer options for the same scenes. Possible randomisation options for *LSF* include randomising the order of scenes, the order of questions asked and the order of possible answer options which the user has to select from. In fact, it had been decided during the design phase of *LSF* not to randomise the questions, so that the trainee should first identify the hazard correctly and then answer questions in an appropriate sequence on dealing with the hazard. However, although the questions are not randomised, the scenes are randomised by the *LSF* back-end system, as well as the sequence in which the options appear.

The highest average for this category (4.7) was allocated to Evaluation Statement 5.2: 'The system monitors and displays student progress'. This is due to *LSF* clearly displaying at the top of the screen the position where the trainee is in terms of the scenes, as well as the current score.

8.4.1.5. *LSF* heuristic evaluation overall results

The heuristic evaluation instrument contains a total of 105 evaluation statements. When taking a holistic view over all the data presented from all four categories, the average rating for each evaluator was calculated as the total of the scores allocated by the evaluator divided by the number of evaluation statements. This yielded the following results, as shown in Table 8.7:

Table 8.7: Overall averages per evaluator for the LSF prototype.

Evaluator	A	B	C	D	E	F	All
Average	4.4	3.7	4.4	3.8	3.2	3.6	3.9

As indicated in Table 8.7, the overall average rating of all six evaluators for the complete evaluation is 3.9 out of 5, which is equivalent to 76%. Furthermore, from Table 8.7 it is clear that there is little difference between the ratings allocated by the usability experts, Evaluators A and B (4.4 and 3.7), and the mine training experts, Evaluators C and D (4.4 and 3.8), but the average ratings allocated by the VR experts, Evaluators E and F (3.2 and 3.6) are notably lower.

Table 8.8 summarises the overall average ratings of *LSF* per category. From the four categories, Category 1 (Instructional Design) has the highest average, namely 4.2 (84%) while Category 3 (VR System Design) has the lowest average, namely 3.2 (64%).

Table 8.8: Overall averages per category for the LSF prototype.

Category	1	2	3	4	All
Average	4.2	4.1	3.2	3.9	3.9

Section 8.6.1 (Comparison of heuristic evaluation results) provides further discussion on the category ratings.

8.4.2. Findings of the user satisfaction survey of the *LSF* prototype

The user satisfaction questionnaire was completed by 195 participants, who were mine workers returning from their annual leave, so called *ex-leaves.* They underwent training with the *LSF* prototype as a refresher course prior to working underground again. The next three sections discuss the findings of the survey with regard to the biographical details of the participants; data regarding the *LSF* prototype; and the open-ended section on user comments, respectively. Interpretation of the findings is given where the findings are presented.

8.4.2.1. Biographical information

The 195 participants represented a total of 23 different job titles as indicated on the survey forms. Table 8.9 lists the 23 job titles, the number of participants holding each and the percentage each number represents. These job titles are representative of a large spectrum of underground mine workers, with the most participants being winch operators (51), equipment helpers (35), loco operators (20) and rock drill operators (19).

Table 8.10 indicates the school education level completed by the participants, showing that 57.4% of them had completed matric (Grade 12 level). This means that just over 40% of the participants had left school without matric, of whom nine participants had only completed primary school (Grade 6 or 7), while five participants did not respond to this question.

Table 8.9: Participant job titles, number and percentage of incumbents.

Job Title	Number	Percentage
Chairlift operator	1	0.5
Construction team leader	5	2.5
Conveyor belt attendant	3	1.5
Drill rig operator	3	1.5
Engineering assistant	4	2.1
Equipment helper	35	17.9
Grade control observer	3	1.5
Gunnite helper	3	1.5
Health and safety official	2	1
Instructor	1	0.5
LHD operator	2	1
Loco operator	20	10.3
Panel operator	9	4.6
Rigger helper	9	4.6
Rock breaker operator	3	1.5
Rock drill operator	19	9.7
Sampling assistant	4	2.1
Sectional gang leader	2	1
Store issuer	7	3.6
Survey crew leader	1	0.5
Survey helper	7	3.6
Vacuum pump operator	1	0.5
Winch operator	51	26.2
Total	195	100

Table 8.10: LSF participants' schooling levels.

School level completed	Number	Percentage
< Grade 6	0	0
Grade 6 – 7	9	4.6
Grade 8 – 9	15	7.7
Grade 10 – 11	54	27.7
Grade 12	112	57.4
No response	5	2.6
Total	**195**	**100**

In terms of exposure to technology, Table 8.11 shows how many participants had previously used computers, cell phones and bank ATMs. Of particular relevance is the fact that only 66 (33.8%) of the 195 participants had prior exposure to computers before doing the *LSF* training, but their exposure to cell phones (95.9%) and bank ATMs (88.7%) was considerably higher.

Table 8.11: LSF participant exposure to technological devices.

Device used before	YES	YES %	NO	NO %
Computer	66	33.8	129	66.2
Cell phone	187	95.9	8	4.1
Bank ATM	173	88.7	22	11.3

Table 8.12 shows that 166 of the 195 participants did the voluntary pre-training on how to use the computer mouse before doing the *LSF* prototype. It is therefore important to note that, even though 33.8% had used computers before, 85.1% of participants had opted to do the mouse pre-training before doing the *LSF* training. This indicates that they were not comfortable with a mouse, and that in many cases the previous computer exposure indicated in Table 8.11 was probably minimal.

Table 8.12: LSF participants choosing to do pre-training.

Pre-training done?	Number	Percentage
YES	166	85.1
NO	29	14.9
Total	195	100

8.4.2.2. Data regarding the *LSF* prototype

As discussed in Chapter Four, Section 4.4, effective communication is a challenge as the workforce of the mining industry speaks a range of languages. Moreover, lack of literacy makes written communication to a large portion of the workforce very difficult. Table 8.13 indicates the home language of the participants as well as the language they selected when they commenced their sessions on the *LSF* prototype, with regard to the four languages for which *LSF* makes provision.

Table 8.13: Participants' home languages and selected LSF language.

Home Language	Number	%	*LSF* selected language	Number	%
English	2	1	**English**	42	21.5
Xhosa	21	10.8	**Xhosa**	19	9.8
Tswana	122	62.6	**Tswana**	132	67.7
Sepedi	11	5.6	**Sepedi**	2	1
Other	39	20	-	-	-
TOTALS	**195**	**100**		**195**	**100**

From the information presented in Table 8.13 it is evident that:

- The four languages available in the *LSF* prototype apply to the home language of 156 of the participants (80%), with the remaining 39 (20%) having a different home language not available in the prototype. These participants had to choose one of the four available languages, which explains why more participants chose English and Tswana for the *LSF* prototype than those who indicated these languages as their home languages. This is particularly true for English, which was the home language of only two participants, yet 42 selected it as their *LSF* language.
- Some Sepedi and Xhosa speakers actually chose to do the prototype in a different language, even though Sepedi and Xhosa were available to them. This could be due to the fact that English is the language mostly used for instruction at the mine and they might have felt more comfortable undergoing training with mining terminology in English. Furthermore, the Rustenburg area, where the survey was conducted, is predominantly a Tswana-speaking region, and many mine workers may be more exposed to the Tswana terminology in their daily work environment.

The user satisfaction questionnaire (Appendix B-2), though based on relevant theory, was not structured according to the *DEVREF* Framework (discussed in Section 5.8), as was the case with the heuristic evaluation framework. For the presentation of the questionnaire findings related to the *LSF* prototype (Questions 2.2 to 2.20 of the questionnaire), the findings were grouped under various usability-related categories: Ease of use (Rogers *et al*., 2011), Learnability (Rogers *et al*., 2011), Satisfaction (Masemola & De Villiers, 2006), Authenticity (Amiel & Reeves, 2008; Marchand & Walker, 2009), and Method of choice (Alessi & Trollip, 2001). The question numbers from the questionnaire relating to each category are as follows:

- Ease of use: 2.4, 2.5, 2.6, 2.7, 2.10 and 2.14;
- Learnability: 2.9 and 2.15;
- Satisfaction: 2.2, 2.3 and 2.16;
- Authenticity: 2.8, 2.11, 2.12, 2.13 and 2.17; and
- Method of choice: 2.18, 2.19 and 2.20.

The categories above were not indicated in the user satisfaction questionnaire and the questions related to specific features of usability were deliberately presented in no particular order. It should also be noted that the terminology used for Likert scaling varied according to the particular question, for example, the scale used for the question 'How easy was it to work with the mouse?' was *Very easy*, *Easy*, *Average*, *Difficult* to *Very difficult*, whereas for the question 'How realistic were the accidents that you saw in this training program?' the participant could select between *Very realistic*, *Realistic*, *Average*, *Not really realistic* and *Not at all realistic*.

Although the questions were not phrased in standard terminology, the distribution over the five options remained standard in a format of five ratings, basically *Very much, Much, Average, Not much* and *Not at all*. In order to tabulate the findings, this format will be represented in Tables 8.14 to 8.20 as Likert scale options 1, 2, 3, 4 and 5 respectively, ranging from the most positive to the most negative. The only exceptions are Questions 2.11 and 2.12, where a four-option scale was used, and Question 2.20 where a specific preference was selected. Please note that this rating system, where 1 is the most positive option and 5 the most negative, is the opposite of the rating system used for the heuristic evaluation, where the options ranged from *Strongly disagree* as a rating of 1 to *Strongly agree* as rating 5. Consequently, where a rating of 4 to 5 would be the preferred response in the heuristic evaluation, a rating of 1 to 2 would be positive feedback for the user satisfaction survey.

The next subsections present the findings per category, presented as percentages, rather than the actual number of participants. In some cases, not all of the 195 participants responded to all the questions. Such responses are indicated in a separate column as NR (No Response). The last column in each table is not a percentage value, but indicates the average Likert scale rating of all the participants for each question and for the category.

Ease of use

The responses received for Questions 2.4, 2.5, 2.6, 2.7, 2.10 and 2.14 are grouped together in an ease-of-use category. Table 8.14 summarises the findings for the questions related to the ease of use of *LSF.*

Table 8.14: Responses for the Ease of use category – LSF prototype.

Rating:	1	2	3	4	5	NR	AVE
2.4. How easy was this training program to use?	44.6	40.5	11.8	1.5	0	1.5	1.7
2.5. How easy was it to work with the mouse?	69.3	24.6	3.6	1	0.5	1	1.4
2.6. How much assistance did you require from the facilitator?	42.5	13.3	3.6	14.9	23.1	2.6	2.6
2.7. How well could you understand the questions in the program?	55.4	37.4	5.6	0.5	1	0	1.5
2.10. How much are you at ease using computers for training?	64.1	24.6	5.6	3	2.6	0	1.6
2.14. Were you given enough time to complete the training program?	64.6	22.6	4.1	0.5	7.2	1	1.6
Average:	**56.8**	**27.2**	**5.7**	**3.6**	**5.7**	**1**	**1.7**

It is clear from Table 8.14 that the majority of responses received were in the first and second scale options, representing *Very much* and *Much* respectively. Question 2.6 deals with the amount of assistance required by the users. In the way this question was phrased in the questionnaire, the desired response would be rating 4 or 5, demonstrating that the user did not require much assistance in using the system. For the purpose of consistency, the results for Question 2.6 are presented in the reverse order in Table 8.14. This means that the participants' 5 responses will now be indicated in the 1 column, the 4 responses in the 2 column, the 2 responses in the 4 column, the 1 responses in the 5 column, and the 3 responses remain in column 3. This way the most positive responses are indicated first, which is in line with the order of all the other responses, and uniform data is provided for statistical analysis.

The results for Question 2.6, showing a 2.6 average, are not surprising, due to the fact that more than 85% of these users opted to do the pre-training system, as shown in Table 8.12, indicating that they were not regular computer users and needed assistance. This could also explain why some users indicated in Question 2.14 that they did not have enough time to complete the system. They may, in fact, have been struggling with general use of the computer. In assessing the responses to Questions 2.4, 2.5, 2.7 and 2.10, the total percentages in the first and second scale options accumulatively for these four questions are 85.1%, 93.9%, 92.8% and 88.7% respectively, which represents an average of 90.1% positive responses for the 195 users.

Learnability

Table 8.15 reports a majority of positive responses relating to the learning value of the system, with 93.3% and 96.9% of the responses being in the first and second rating options for Question 2.9 and Question 2.15 respectively. Of these positive responses, the greater percentage of responses (73.6% on average) was for the first option, and 21.3% for the second option.

Table 8.15: Responses for the Learnability category – LSF prototype.

	Rating:	**1**	**2**	**3**	**4**	**5**	**NR**	**AVE**
2.9.	**How much did you learn by using this program?**	64.1	28.7	4.6	0.5	1.5	0.5	1.5
2.15.	**Will this training program help you to be more aware of the hazards in the workplace?**	83.1	13.8	2.5	0	0	0.5	1.2
	Average:	**73.6**	**21.3**	**3.6**	**0.3**	**0.8**	**0.5**	**1.4**

Satisfaction

Table 8.16 presents responses to the *Satisfaction* category, comprising questions related to enjoyment, interest and satisfaction.

Table 8.16: Responses for the Satisfaction category – LSF prototype.

		Rating:	1	2	3	4	5	NR	AVE
2.2.	**How interesting was this training program to you?**		72.3	21	2.6	1	0.5	2.6	1.3
2.3.	**How much did you enjoy doing this program on the computer?**		73.3	21	3.1	1.5	0	1	1.3
2.16.	**How satisfied are you with the feedback that you received from the program while you were doing the training?**		62.1	32.4	3	1	1.5	0.5	1.5
		Average:	**69.2**	**24.8**	**2.9**	**1.2**	**0.7**	**1.37**	**1.4**

Once again, nearly all the responses fell in the first two rating options, with most of them in option 1. For Question 2.2, 146 responses fell under option 1, *Very interesting*, representing 72.3% of the participants. For Question 2.3, 73.3% of the participants enjoyed the system *Very much*, while in response to Question 2.16, 62.1% were *Very satisfied* with the system feedback. This gives an average of 69.2% in the first option for the *Satisfaction* category. None of the questions yielded sufficient negative responses to warrant further analysis.

Authenticity

Five questions are grouped together in the *Authenticity* category: Questions 2.8, 2.11, 2.12, 2.13 and 2.17. The responses to Questions 2.8, 2.13 and 2.17 are summarised in Table 8.17, and the responses to Questions 2.11 and 2.12 in Table 8.18. The reason for separating Questions 2.11 and 2.12 in the analysis, is due to the way they were asked in the questionnaire. A four-option scale was used, as the nature of the questions did not make an *Average* option feasible. Questions 2.11 and 2.12 dealt with the occurrence of the accidents shown in the system and the options used were *None of them*, *Some of them, Most of them* and *All of them*.

Table 8.17: Responses to Questions 2.8, 2.13 and 2.17 of the Authenticity category – LSF prototype.

	Rating:	1	2	3	4	5	NR	AVE
2.8.	**How easily did you recognise the objects on the screen?**	45.4	42.3	9.3	2.6	0.5	0	1.2
2.13.	**How realistic were the accidents that you saw in this training program?**	48.2	39.9	7.3	2.1	2.6	0	1.7
2.17.	**To what extent are the geological hazards shown in this program relevant to your job?**	46.7	35.4	11.3	4.6	1.5	0.5	1.8
	Average:	**46.8**	**39.2**	**9.3**	**3.1**	**1.5**	**0.2**	**1.6**

The responses listed in Table 8.17 follow a similar trend to the previous categories, where most of the responses were option 1 or 2. In terms of the recognisability of the graphic objects used in the system (Question 2.8), 87.7% of the responses were in scale option 1 or 2. Similarly, 88.1% of the participants selected rating option 1 or 2 to assess the realism of the accidents (Question 2.13), and 82.1% indicated likewise for the relevancy of the hazards (Question 2.17). However, the occurrence of option 1 was, on average, lower (46.8%) than in the responses reported in Tables 8.14 (56.8%), 8.15 (73.6%) and 8.16 (69.2%).

Table 8.18: Responses to Questions 2.11 and 2.12 of the Authenticity category– LSF prototype.

	Options:	None	Some	Most	All	NR
2.11.	**Do you believe that the accidents you saw in the program can really happen?**	10.3	21	17.4	49.7	1.5
2.12.	**Do you believe that the accidents you saw in the program can really happen to you?**	44.6	30.3	8.7	15.4	1
	Average:	**27.1**	**25.7**	**13.1**	**32.3**	**1.3**

The responses received to Questions 2.11 and 2.12, as shown in Table 8.18, warrant special attention. Even though 67.1% of the participants agreed that *Most* (17.4%) or *All* (49.7%) of the accidents shown could really happen, only 24.1% of them believed that *Most* (8.7%) or *All* (15.4%) of these accidents could actually happen to them. Possible explanations could be that these workers were confident that they work safely and would recognise and address such hazards before an accident occurred. However, it is also possible that some participants believed that accidents would not happen to them. This is also evident from the twenty responses (10.3%) to Question 2.11 which suggested that none of these accidents can actually happen at the mine.

This is of particular concern, since the accidents selected for the system were generic in nature and occur frequently in the mining industry. In fact, most of them were based on actual underground incidents at the very mine where these participants were employed.

Method of choice

This category deals with questions related to the computer as training medium. Table 8.19 shows the responses to Questions 2.18 and 2.19. In response to Question 2.18, a total of 183 participants (93.9%) indicated that they would *Very much* or *Much* like to do similar training on the computer again. Only two participants (1%) indicated that they did not want to do similar computer-based training again. Similarly, in response to Question 2.19, 180 participants (92.3%) selected scale option 1 or 2 to indicate their preference for this type of training instead of classroom training only.

Table 8.19: Responses to Questions 2.18 and 2.19 for the Method of choice category – LSF prototype.

	Rating:	**1**	**2**	**3**	**4**	**5**	**AVE**
2.18.	**How much would you like to do training like this on the computer again?**	71.8	22.1	5.1	1	0	1.4
2.19.	**Do you think this type of training on the computer is better than just listening to an instructor in the classroom?**	76.4	15.9	7.2	0	0.5	1.3
	Average:	**74.1**	**19**	**6.2**	**0.5**	**0.3**	**1.4**

Table 8.20 lists the responses to Question 2.20, which dealt with participants' preferred method of training. The options listed were classroom lectures, practical training, video-based training, computer-based training and a combination of lecture and computer-based training. In hindsight, the *Practical* training option should not have been included in this questionnaire, since the intention of this research is to promote the use of computer-based training as supplementary to the practical training, and not as a replacement. Due to the nature of these jobs, hands-on training remains essential in the mining environment to ensure that employees practise the skills prior to working in such hazardous conditions. Nevertheless, the majority of participants still selected computer-

based training (31.8%) or the combination of lectures and computer-based training (43.6%) as their preferred training option, that is, a total of 75.4% preferred training via the computer, although more than half of them chose to do it in combination with lectures in a classroom situation. The very low support for video-based training as the preferred method is notable – three participants only, which is 1.5%.

Table 8.20: Responses to Question 2.20 for the Method of choice category – LSF prototype.

Options:	Lecture	Practical	Video	Computer	Lecture and Computer	NR
2.20. Please indicate your preferred method of training	7.2	14.4	1.5	31.8	43.6	2

It is evident from the responses shown in Tables 8.14 to 8.20 that there is great support for training via the *LSF* approach. The prototype obtained high ratings in all the categories, particularly in learnability and satisfaction.

The next section deals with the qualitative comments received via the open-ended questions in the questionnaire.

8.4.2.3. Qualitative user comments on the *LSF* prototype

This section analyses the responses received to the five open-ended questions, and provides selected quotes. Interpretations and discussions of the responses are interspersed between the findings. The third part of the user satisfaction questionnaire contained five open-ended questions:

1. What do you think are the best features of the program?
2. What aspects of the program do you think should be improved?
3. Is there any other training that you would prefer to do on the computer?
4. Please describe problems you encountered in using the *Look, Stop and Fix* program.
5. Do you have any other comments on the *Look, Stop and Fix* program?

Of the 195 participants, only 41 responded to the open-ended questions and, in many cases, responses were received only for some of the questions. This section analyses qualitative comments, identifies themes and patterns, and is illustrated by selected quotes.

Most responses to Question 1 regarding the best features of *LSF* indicated that the program improves knowledge and understanding of how accidents happen, thereby increasing awareness of dangers. Several positive comments were received regarding the individualised nature of the training system and its user friendliness. Other specific comments received from different participants, were:

- 'It is a good induction for workers, showing the underground environment'.
- 'The program shows the importance of working according to the set standards'.
- 'You must follow the rules'.
- 'It teaches you to be alert'.
- 'It reminds you of things you may have forgotten'.
- 'The program is easy to follow'.

Diverse suggestions were received for possible improvements to the system, as requested by Question 2. In categorising the answers, the main areas identified for improvement were:

- Content improvement: Some of the content additions requested, were scenes on methane gases, the use of pinch bars, winch operation, hanging wall conditions leading to falls of ground, velocity aspects and hydrometer usage. One comment was 'workers showing up drunk should be added as a hazard'.
- Usability improvement: Four participants requested that more languages should be incorporated. Further comments made by various participants were:
    - 'Please show wrong and correct answers as feedback',
    - 'Some scenes could disturb sensitive viewers',
    - 'Some accidents are a bit shocky',
    - 'Improve graphics e.g. rails and sleepers',
    - 'I struggled with the mouse',
    - 'Everybody should do this three times a year', and
    - 'I would like to see full 3D'.

Question 3 dealt with the identification of other training that could also be delivered via computers. Although fifteen participants answered 'anything', more specific responses were: treating injured workers, health education, teamwork as a gang, strata control, locomotive training, explosives handling, trackless mobile equipment, installation of pipes, winch operation and conveyors.

Question 4 requested that the participants mention any problems experienced while using the *LSF* system. As with Question 2, two kinds of answers were received, some relating to content and others to usability. The content-related answers referred to aspects in scenes that were not according to mine standards, but were not the specific hazard required to be identified by the system. These included:

- ‘They are working at height without safety harnesses’,
- ‘Working on railway line without putting stop signs’,
- ‘Winches are not barricaded’,
- ‘Did not understand persons fainting due to poor ventilation’, and
- ‘I gave the right answer, but the computer did not agree with me’.

Examples of usability-specific comments received, were:

- ‘Voice is not corresponding with text’,
- ‘I had a problem understanding some of the words in Tswana’, and
- ‘… struggled to change volume’.

Question 5 required that participants write down any other comments about the system. Only five comments were received:

- ‘We don’t fix the workplace, just start the job’,
- ‘This should be done annually’,
- ‘I love it, think it is a brilliant idea’,
- ‘Program is very wonderful, very refreshing, keep it up!’, and
- ‘It’s a good program, well done to training centre’.

8.4.2.4. Discussion and interpretation of findings of LSF

The responses to the usability-related categories Ease of use, Learnability, Satisfaction and Authenticity were presented in Tables 8.14, 8.15, 8.16 and 8.17, respectively. As indicated in these tables, the overall average ratings for these categories were:

- Ease of use: 1.7;
- Learnability: 1.4;
- Satisfaction: 1.4; and
- Authenticity: 1.6.

According to the rating system used, a 1 rating is the most positive rating. It is therefore clear that the participants rated Learnability and Satisfaction as the best usability traits, slightly better than Authenticity, which is followed by Ease of use. Although the average ratings of Learnability and Satisfaction are both 1.4, from Tables 8.15 and 8.16 it is clear that participants allocated more 1 ratings to Learnability (73.6%) than to Satisfaction (69.2%). Regardless of this comparison, all four averages are between a 1 and 2 rating, meaning that all of these usability traits were evaluated as between very positive and positive.

The comments received in the open-ended section, indicated high praise for the system, but also identified several areas for improvement, in terms of both content and usability.

With regard to content, the comments received for Questions 1 and 5, indicated the success of the *LSF* prototype in increasing awareness of hazards and in informing trainees of the possible consequences of not addressing these hazards correctly. There were also calls to have such training on a regular basis and not only once. After the success of the *LSF* prototype system, management at the mine where the evaluations were done, decided to implement this system as compulsory for ex-leaves on an annual basis.

In responding to Question 2, seven participants suggested that additional hazards should be incorporated. The purpose of the prototype had been just to select a number of generic hazards in order to evaluate the concept, rather than to present a comprehensive set covering all possible hazards. The responses do, however, confirm the concept and the need to expand this type of training system from generic hazards to other areas of the workplace where job-specific hazards may occur. This notion is also supported by participants suggesting other possible training systems that can be developed using virtual reality, as indicated by the answers received for Question 3.

The content-related aspects mentioned in the answers to Question 4 do indeed warrant consideration. The possibility of other hazards or substandard conditions being present in a scene intended to portray one specific hazard, can be a confound and can have negative consequences:

- A trainee could spot a condition that implies the presence of a hazard in the scene and then continue to the identification step where a selection must be made between five possible hazard names displayed as labels on buttons. If the scene was actually intended for the identification of another hazard, and not the one spotted by the trainee, then he might select an incorrect answer, marks would be deducted, and it would go on record that the trainee had identified the hazard incorrectly. If another hazard was indeed portrayed in the scene, then this would be unfair towards the trainee.
- In cases where the hazards presented as options do not contain the other hazard that the trainee spotted, the trainee might infer that the hazard he thought he

saw, was not a hazard. This could lead him to believe it is acceptable in the workplace.

- This argument would similarly apply to scenarios where a trainee might see more than one hazard, correctly identify the one listed in the selections, and then receive positive feedback from the system. This could lead to the trainee inferring that the other hazards spotted, which were not indicated as such by the training system, might not be hazards and may indeed be acceptable practice.

It thus became clear that the prototype should be refined to prevent confounds and ambiguities. This can be done either by ensuring that there are no other hazards in scenes except those intended to be spotted, or by clearly indicating to the user that there may be more than one hazard in a scene, but that they should select an answer from the given list. Moreover, it should be ensured that the hazard names in the selection list do not include hazards that might indeed be present as additional hazards.

Some of the usability improvements mentioned in response to Question 2 and Question 4, can be accommodated in an upgrade of the prototype. Some comments were vague, for example, 'I gave the right answer, but the computer did not agree with me' fails to indicate whether this was a program error or a knowledge error on the part of the participant. Since this comment was only received once, it may be the latter.

The comment regarding not understanding the Tswana terminology does warrant further investigation. The prototype was initially developed in English and was translated into Tswana, Sepedi and Xhosa once all the text and narration had been established. The translations were done by a professional translation service. A potential problem is that translators may not be familiar with the mining environment and its terminology, which could result in a 'pure' translation, yet one that uses terminology other than the terms actually used on site. An example is the term 'near-miss incident', which within a safety environment indicates an incident that could have resulted in an injury or damage to property. The English terms, 'near-miss incident' or simply a 'near miss', are used on site regardless of the language of the conversation. Should a translator literally translate the English words 'near', 'miss' and 'incident' into their correct Tswana words, it would result in a phrase not understood at all, not even by Tswana speakers.

Furthermore, the researcher is aware that there are different dialects of Tswana and Sepedi used in different parts of South Africa. For example, a Sepedi translation used successfully in the Steelpoort environment in the Limpopo Province would not be completely acceptable in the Rustenburg environment in the North West Province. One way to ensure that a translated text is indeed usable within a particular environment in the mining industry would be to have the text translated by a person from that particular site, who is familiar both with English and the translated language. Alternatively, a translated text should be checked by a site expert prior to implementing it.

From the analysis of the user satisfaction questionnaire, it is clear that it was successful as an instrument to determine important user information, to evaluate user satisfaction, and to establish the strengths and weaknesses of the prototype.

Nevertheless, there were certain weaknesses in the questionnaire, namely:

- Allowing the participant to write his job title resulted in many different versions of the same job, for example, the job title Equipment Helper, was indicated by certain participants as Equipping Helper, Equipment Assistant, Equip Help, Equipping or Production Equipment Helper. There were also various versions of the spelling. This resulted in difficulties in coding the data. The 23 major job titles should have been provided in a list from which participants could have selected the relevant one.
- Question 1 in the open-ended questions section required the participants to list the 'best features of the program'. The responses showed that many participants did not understand the term 'features of the program'. This should have been better explained.

8.4.2.5. Internal consistency reliability: *LSF* questionnaire responses

The discussion of the results of Table 8.14 in Section 8.4.2.2 focused on the concept of ease-of-use of a training application. Therefore, in a further analysis step, the possibility of a single measure of ease-of-use perception (instead of separate response measures on each question pertaining to ease of use) was examined. It was argued that if a single (thus parsimonious), and reliable measure of ease-of-use perception of training could be derived, the possible effect of biographical attributes of participants on their perceptions of how easy it is to train with a specific application, could be identified. Once identified, such peripheral issues that could hamper training (e.g. educational level or computer training, or training to improve skills of handling a mouse) can be addressed.

In a more in-depth analysis, the reliability (the internal consistency reliability) of such a measure – and the most appropriate subset of questionnaire questions to describe the concept of ease of use of the training application – was investigated by means of a scale reliability test. The measure for internal consistency reliability, namely, Cronbach alpha, is based on the inter-correlation between variable ratings (responses to questionnaire items). If the inter-correlation is high, as measured by the Cronbach alpha test, then it serves as an indication that the individual items all contribute towards measuring the same concept. The theoretical value of alpha varies between 0 and 1.The larger the value of Cronbach alpha, the more desirable (Ritter, 2010). A value in the region of 0.7, or greater than 0.7, is regarded as a good indicator of internal consistency reliability.

A prerequisite for scale reliability testing is the availability of at least four sets of responses to four questions of the questionnaire under investigation (George & Mallery, 2003; Kline, 2000). This explains why only the *Ease of use* category in the questionnaire (with a possible subset of six questions) was investigated for internal consistency reliability on the concept of ease of use of the training application and the calculation of a parsimonious perception measure. As discussed in the Section 8.4.2.2, the *Ease of use* category consists of the following questions:

2.4: How easy was this training program to use?
2.5: How easy was it to work with the mouse?
2.6: How much assistance did you require from the facilitator?
2.7: How well could you understand the questions in the program?
2.10: How much are you at ease using computers for training?
2.14: Were you given enough time to complete the training program?

Initial scale reliability tests indicated that Question 2.6 and Question 2.14 did not contribute towards explaining a parsimonious ease-of-use perception measure. Consequently, the underlying implications of these two questions were further investigated: Question 2.6 was phrased in such a way that responses to the question could suggest other factors than the use of the program itself, for example, computers not working properly or network issues. Similarly, responses to Question 2.14 could be negative due to facilitators not allowing sufficient time for trainees to complete using the *LSF* system, and not necessarily due to trainees struggling to use the system. Therefore it was decided to omit the responses to Questions 2.6 and 2.14 from the final consistency reliability test that is reported in the next subsection and only to evaluate how well the remaining four sets of responses (namely the responses to Questions 2.4, 2.5, 2.7 and 2.10) describe each participant's perception of the ease of using the *LSF* training system.

The two biographical properties mentioned in the introductory argumentation of this section – to be investigated for their possible impact of perceptions of ease of use of training – are reported in Tables 8.10 and 8.12 respectively. Table 8.10 showed the responses received for participants' schooling levels, based on the responses received for Question 1.1: What was the highest standard or grade that you completed at school? Table 8.12 indicated the responses to Question 1.3: Did you do the PC literacy pre-training on how to use the computer mouse before you started the *Look, Stop and Fix* program? If the reliability of a parsimonious ease-of-use measure can be verified, the scores for each participant on such a measure can be calculated and a further analysis can be performed on this set of perception scores to evaluate the significance of the effect of the two mentioned biographical attributes on ease of use of training perception.

The next subsections present the results of the internal consistency reliability verification, the calculation of the ease-of-use scores and the results of an analysis of variance to determine the statistical significance of the effect of schooling levels and pre-training on perceptions of the ease of use of training.

Internal consistency reliability

A final scale reliability test was performed on the four sets of responses to Questions 2.4, 2.5, 2.7 and 2.10 and yielded a Cronbach alpha value of 0.64. This was considered an acceptable indication of internal consistency, especially viewed against the fact that the purpose of the research (and measure) was investigative. With internal consistency reliability verified, the analysis could proceed to calculate an ease-of-use score on the questionnaire item responses by calculating a mean rating score for each participant based on the rating scores for the separate questions for each participant.

The calculation of the ease-of-use score

Because this analysis investigates the possible effect of *schooling level* and *prior PC training* on participants' perceptions of the ease of use of *LSF*, a summary to illustrate the calculated ease-of-use perceptions scores is presented in Table 8.21. Ease-of-use mean scores were calculated for the categories of *schooling level* (Grades 6-7; 8-9; 10-11; and 12) for Question 1.1 and *pre-training* ('yes' and 'no') for Question 1.3. This summary thus provides a first suggestion of the nature of the possible effect of prior training and schooling on ease-of-use perceptions.

Table 8.21 reports the number of responses for each of the schooling-by-pre-training combination levels, the mean ease-of-use perception score and standard deviation for these level combinations. For example, the first row in the table indicates that the mean perception score for participants with a Grade 6-7 schooling level and prior training exposure (seven in total) is 1.68; and the fourth row indicates that for participants with a Grade 8-9 schooling level and no prior training exposure, the mean perception score is 2.25.

Since the ease-of-use score is interpreted similarly to the rating values of the Likert rating scale of the user satisfaction questionnaire, a mean ease-of-use score of 3.5 – which can be interpreted as a 'somewhat difficult' rating – reported for Grade 6-7 schooling level and no pre-training group (highlighted in blue) suggests a different perception from a mean score of, for example, 1.43 – which reflects a perception of 'very easy' to 'easy' - for the Grade 12 group who did complete the pre-training (highlighted in yellow). Such suggestions are firstly investigated and then verified in the next two sets of analyses to follow. No schooling level less than Grade 6-7 was reported, therefore the first schooling level, level 1, was omitted in reporting (although indicated in Table 8.10).

Table 8.21: Means table for the ease-of-use analysis for LSF responses.

Schooling	Pre-training?	Mean	Standard deviation	Number
Grade 6-7	Yes	1.68	0.57	7
	No	3.50	0	2
Grade 8-9	Yes	1.62	0.46	13
	No	2.25	0.71	2
Grade 10-11	Yes	1.61	0.70	45
	No	1.72	0.65	9
Grade 12	Yes	1.43	0.43	94
	No	1.74	0.63	18

An initial observation from Table 8.21 is that, for all four schooling levels, the mean scores of participants who completed the pre-training are lower than those who did not. This suggests that those who underwent pre-training, found the system easier to use than those who did not, regardless of schooling level (a score of 1 represents 'very easy' and a score of 5 'very difficult').

It is also clear from the table that the largest group of participants (94 of 195 = 48%) had a Grade 12 schooling level (i.e. matric) and did complete the pre-training system. Furthermore, it is notable that this group had the lowest mean score (1.43, which is between 'very easy' and 'easy'), suggesting they found the program the easiest to use. These observations are statistically verified in the next subsection.

General linear model approach to analysis of variance

In this subsection the results of an analysis of variance on the ease-of-use perception scores for participants evaluated against the effect of *schooling* and *pre-training,* are discussed. This section serves to validate the statistical significance of trends suggested in the previous subsection. The effects of schooling level and pre-training are included in the analysis of variance model to evaluate the statistical significance of these effects on perceptions of how easy it is to use *LSF*. The general linear model (GLM) approach to analysis of variance was used to accommodate the fact that the four categories of schooling levels and pre-training were not represented by equal numbers of participants, thus producing unbalanced data.

This analysis is presented in Table 8.22, in four separate blocks:

- Block 1 indicates that 195 observations were used, and gives the mean ease-of-use rating score the dependent variable) for all these participants, the standard deviation of this score and the minimum and maximum rating values.
- Block 2 indicates the levels of the independent effects included in the model, representing the independent variables (schooling has four levels; and pre-training has two levels).
- Block 3 indicates that the responses of 190 participants were included in the analysis (some participants did not respond to all four questions that constitute the ease-of-use score and were excluded from the analysis: the 'missing' observations).
- Block 4 reports on the analysis of variance itself, indicating in row two that the general F statistic of 5.37 for the test is statistically significant on the 1% level of significance. Furthermore, in rows three to five, in the last two columns it is indicated that the effect of schooling, pre-training and schooling-training interaction is statistically significant on the 1% level of significance (the probabilities (Pr) of 0.002; <0.0001 and 0.0048, associated with the respective F statistics of 6.93; 19.75 and 4.45, are all <0.01, which signifies statistical significance on the 1% level of significance).

Table 8.22: GLM analysis of variance for ease of use for LSF responses.

N	Mean	Standard Deviation	Minimum	Maximum
195	1.5576923	0.5812290	1.0000000	5.0000000

Class	Levels	Values
Schooling	4	2 3 4 5
Pre-training	2	1 2

Number of Observations Read	195
Number of Observations Used	190

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	11.099	1.586	5.37	<0.0001
Schooling	3	6.135	2.045	6.93	0.0002
Pre-training	1	5.829	5.829	19.75	<0.0001
Schooling*Pre-training	3	3.936	1.312	4.45	0.0048
Error	181	53.419	0.295		
Corrected Total	188	64.519			

It can thus be deduced that the analysis of variance results statistically verify the suggestion of Table 8.21 that schooling level, pre-training, and their interaction affect perceptions of ease of use of the computer training in a way that is statistically significant.

Bonferroni multiple comparisons of means tests

Analysis of variance results indicate that schooling and pre-training *affects* perceptions of how easy it is to use the training application, but does not indicate *how* perceptions are affected. Bonferroni multiple comparisons of means tests were therefore conducted to verify the nature of the effect of schooling and pre-training on ease-of-use perceptions.

Table 8.23 describes the nature of these effects: mean perceptions scores for both the schooling and pre-training effects are compared (separately for each effect) on a pair-wise basis in a multiple t-test (with the significance level adjusted to accommodate multiple pair-wise comparisons for schooling levels), to test which pairs of means differ statistically significantly from one another. The means that differ statistically significantly from one another have different small letters associated with their means (indicated as the Bonferroni grouping indicator in Table 8.23). Higher schooling levels contribute to the ease with which the training application is used: Grade 6-7 mean perception levels is statistically significantly higher ('more difficult') than Grade 12 perception means, in that a mean of 2.08 (Grade 6-7) differs statistically significantly from a mean of 1.48 (Grade 12).

The same applies to the interpretation of prior and no prior training means of 1.88 (approximately 2, which is 'easy') and 1.50 (between 1 and 2 which would mean 'halfway between very easy and easy'). The least significant difference (lsd) 5% significance-indicators used in these tests are also reported in Table 8.23.

Table 8.23: Bonferroni multiple comparisons of means tests illustrating effect of schooling and pre-training on ease-of-use perceptions.

Means with the same letter are not significantly different.			
Bonferroni Grouping (lsd = 0.46)	**Mean**	**N**	**Schooling**
A	2.08	9	Grade 6-7
Ab	1.70	15	Grade 8-9
Ab	1.63	54	Grade 10-11
B	1.48	111	Grade 12
Bonferroni grouping (lsd = 0.21)			**Pre-training**
B	1.50	159	Pre- training Yes
A	1.88	30	Pre-training No

Apart from the nature of the main effects discussed in the previous paragraph, Table 8.22 reports a statistically significant interaction effect on perceptions as well. However, the contribution of the main effects was considered very substantial in relation to the contribution of the interaction effect (F-values of 6.93 and 19.75 compared to the 4.45 of the interaction effect) and this analysis therefore only focused on the nature of the main effects.

The results of the statistical analysis in this section indicated that higher schooling levels and pre-training contribute to the ease with which *LSF* is used. This finding should be considered in the development of similar training applications:

- Schooling level: As indicated in Table 4.3 in Section 4.4, the schooling level of entrants into the mining industry is increasing drastically. Many local mines now only employ applicants with a matriculation certificate, although there are still many current incumbents with low schooling levels.

- Pre-training: The specially-developed pre-training system covering basic keyboard and mouse skills, discussed in Section 6.7.1, was available to trainees prior to doing the *LSF* system. The use of the pre-training, however, was optional. It is possible that, due to time constraints or overestimating their abilities, some trainees did not do the pre-training that they should have. Furthermore, as more computer-based training systems are introduced to the mining industry, the lack of computer exposure will gradually diminish.

8.5. Evaluation of the Interactive Simulated Geological Conditions prototype

The *Interactive Simulated Geological Conditions* (*ISGC*) prototype was evaluated using the same instruments as the *LSF* prototype, that is, the *DEVREF* heuristic evaluation instrument described in Section 8.2 and the user satisfaction questionnaire described in Section 8.3. These evaluations formed part of the research activities of Cycle 3 of the four cycles of the synthesised design-based research model, as described in the Research Design section in Section 5.6.

The *ISGC* prototype is described in detail in Section 7.3 and the program flow is depicted in Figure 7.9. *ISGC* covers 21 different geological conditions, as well as the associated risks and control measures for each condition. Animations are shown of the possible results of ignoring or not correctly addressing the geological conditions.

The heuristic evaluation was done by the same six experts who evaluated the *LSF* prototype. Their areas of expertise are shown in Table 8.2. This ensured consistency and had the additional benefit of the evaluators already being familiar with the evaluation instrument. The same evaluation process was followed, with the evaluators interacting with the prototype prior to, and during, the evaluation.

The user evaluation was done by 52 trainees, who completed the user satisfaction questionnaire after completion of the *ISGC* training. This was done over a period of two months at the training centre of the same large platinum mine where the *LSF* user evaluation was done. Section 8.5.1 describes the findings of the heuristic evaluation of the *ISGC* prototype, and Section 8.5.2 details the findings of the user satisfaction questionnaire.

8.5.1. Findings of the heuristic evaluation of *ISGC*

Similarly to the presentation of the *LSF* findings in Section 8.4.1, the findings of the heuristic evaluation of the *ISGC* are presented in tabular format, detailing the criteria evaluated and each evaluator's response in terms of the rating (1 to 5), with 1 representing *Strongly disagree* and 5 *Strongly agree*. The tables have the same structure as the *LSF* findings, where each table provides a summary of the heuristic evaluation results giving the individual responses of the six evaluators, an additional column containing the average rating of the evaluators for each statement, and additional rows containing the average per evaluator for each criterion and the category average per evaluator. The same colour scheme is used: the two usability experts' responses are indicated in red, the mining experts' in green and the VR experts' in blue.

The next four tables present the results for the four different categories in the framework respectively:

- Table 8.24 summarises the results of Category 1 – Instructional Design heuristics,
- Table 8.25 covers Category 2 – General Usability heuristics,
- Table 8.26 Category 3 – VR System Design heuristics, and
- Table 8.27 Category 4 – Context-specific heuristics.

Each table is followed by analysis of the tabulated data.

8.5.1.1. *ISGC* heuristic evaluation results for Category 1

Table 8.24: Heuristic evaluation results of ISGC for Category 1: Instructional Design Heuristics.

CRITERIA	Evaluator rating						
	Usability Experts		Mining experts		VR experts		
	A	B	C	D	E	F	Ave
1. Clear goals, objectives or outcomes							
1.1. There are clear goals, objectives or outcomes for the training program.	5	5	5	4	3	5	**4.5**
1.2. Clear goals, objectives or outcomes are communicated at the beginning of the training program.	5	5	5	5	2	4	**4.3**
1.3. The outcomes are measurable.	4	5	5	4	4	4	**4.3**
Criterion average per evaluator	4.7	5	5	4.3	3	4.3	**4.4**
2. Instructional assessment							
2.1. The program provides assessment opportunities that are aligned with the objectives or outcomes.	5	5	5	5	4	4	**4.7**

2.2. The assessment opportunities will serve to enhance trainees' performance.	5	5	5	2	3	4	4
Criterion average per evaluator	5	5	5	3.5	3.5	4	4.4
3. Feedback to user responses							
3.1. The training program provides trainees with constructive and supportive feedback on their performance.	4	4	5	4	3	5	4.2
3.2. The feedback is relevant to the training content.	5	4	5	4	3	4	4.2
3.3. The feedback informs the trainee regarding his level of achievement in the training program.	5	5	5	5	4	3	4.5
3.4. The feedback indicates incorrect responses and provides information on the correct responses.	5	5	5	4	2	3	4
Criterion average per evaluator	4.8	4.5	5	4.3	3	3.8	4.2
4. Motivation and creativity							
4.1. The system supports intrinsic motivation by providing challenges to trainees and encouragement when errors are made.	5	4	5	4	2	4	4
4.2. The program captures the trainee's attention early and retains it throughout.	5	4	5	5	3	4	4.3
4.3. This training program increases trainees' confidence by providing them with reasonable opportunities to accomplish the objectives successfully.	4	5	4	5	2	4	4.2
4.4. The program engages trainees by its relevant content.	5	4	5	5	3	4	4.3
4.5. The program engages trainees by its interactivity.	4	4	5	5	4	4	4.3
Criterion average per evaluator	4.6	4.2	4.8	4.8	2.8	4	4.2
5. Differences between individual users							
5.1. The system takes account of linguistic and cultural differences by allowing trainees to select between different languages.	2	3	5	5	3	5	3.8
5.2. The system caters for trainees with different levels of expertise regarding the content.	3	4	5	2	3	2	3.2
5.3. The system caters for trainees with different levels of computer experience.	4	4	5	4	3	3	3.8
Criterion average per evaluator	3.3	3.7	5	3.7	3	3.3	3.6
6. Reduction of extraneous processing in working memory							
6.1. The training program effectively uses signalling to highlight essential issues, such as restating important points, using headings for important points, or stressing them in audio mode.	5	4	5	2	2	4	3.7
6.2. Redundancy is avoided by not presenting unnecessary information.	5	3	5	4	4	4	4.2
6.3. Redundancy and overload are avoided by not reiterating the same material in multiple modes (.e.g. the program presents information using pictures and spoken words, rather than presenting it in pictures, spoken words, and printed words).	5	4	5	4	3	4	4.2
Criterion average per evaluator	5	3.7	5	3.3	3	4	4

7. Fostering of germane cognitive load							
7.1. The training program supports the formation of mental schema by explaining where newly acquired knowledge fits into the bigger picture.	4	4	5	4	3	3	**3.8**
7.2. The system encourages encoding of the training content into long-term memory by presenting questions after each learning segment.	5	4	5	5	3	4	**4.3**
7.3. Sufficient scaffolding support is provided (in the form of hints, prompts and feedback) to help trainees achieve training goals.	5	2	5	2	2	3	**3.2**
7.4. The training program presents narration in a colloquial conversational style.	4	4	5	4	3	3	**3.8**
7.5. The training program prompts trainees to link concrete example information for each problem category to more abstract information.	5	4	5	4	2	3	**3.8**
Criterion average per evaluator	4.6	3.6	5	3.8	2.6	3.2	**3.8**
8. Appropriate intrinsic cognitive load							
8.1. Working through the training program does not cause trainees to split their attention between multiple sources of visual information.	5	4	5	4	4	4	**4.3**
8.2. The program enhances retention by presenting information in learner-paced segments, rather than as a continuous presentation.	4	4	5	5	4	5	**4.5**
8.3. The system effectively supports dual channel processing of simultaneous visual and verbal material.	5	4	4	4	4	4	**4.3**
Criterion average per evaluator	4.7	4	4.7	4.3	4	4.3	**4.4**
Category Average per evaluator	4.5	4.1	5	4.1	3	3.8	
Overall average rating for Category 1							**4.1**

From the results presented in Table 8.24 it is clear that, as was the case with the Category 1 evaluation of *LSF*, not even one of the evaluation statements received the same rating by all the evaluators. The *ISGC* evaluation followed the same pattern as the *LSF* evaluation in that the two VR experts assigned, on average, lower ratings than the other experts. A possible conclusion here is that, being involved in VR systems development on a daily basis, the VR experts expected more of the prototype within the realm of virtual reality. The main issue to keep in mind, however, is that this is a prototype and serves as a proof of concept for mine safety training improvement.

Analysing the results, Criteria 1, 2, 3, 4, 6 and 8 received high average ratings of more than 4 out of 5. The criteria with lower ratings are Criterion 5 and Criterion 7.

- Criterion 5: Differences between individual users.

  The average rating for this criterion is 3.6. This is mainly due to poor ratings allocated to Evaluation Statement 5.2: 'The system caters for trainees with different levels of expertise regarding the content'. Although Evaluator C allocated a 5 rating, the other evaluators allocated 2 to 4. It is of concern that the two mining experts allocated a 5 and 2 respectively, as this statement deals with mining content. A possible reason for this could be a difference in interpretation by the evaluators. The *ISGC* prototype covers basic information regarding the 27 geological hazards, but does not make a specific distinction between novice or experienced users. One possible interpretation of the evaluation statement may be a focus on the level of the content, rather than on the actual intention of the evaluation statement, which is to evaluate whether the system distinguishes content-wise between different levels of users. In a comment received in the open-ended space at the end of the category, Evaluator D, who assigned a 2 rating to this statement, wrote that 'only one level of expertise is available'. Since the *ISGC* prototype does not have different user levels, the higher ratings assigned by other evaluators seem to imply that the basic level of the content should be generic knowledge to all users in the domain and that differentiation in levels of expertise may not necessarily apply to this scenario. However, to prevent misunderstanding, the evaluation statement could be rephrased to express more clearly what should be evaluated.
- Criterion 7: Fostering of germane cognitive load.

  The average rating for this criterion is 3.8, the main reason being the low ratings for Evaluation Statement 7.3: 'Sufficient scaffolding support is provided (in the form of hints, prompts and feedback) to help trainees achieve training goals'. Two evaluators each assigned a 5, but the other ratings are three 2's and one 3. Once again the question arises: why would some evaluators assign a high score and others a low score to the same evaluation statement? In this case, the difference probably arises from the use of the word 'sufficient' in the evaluation statement. What could be sufficient for some evaluators may not be so for others. Another possibility is that, in working through the prototype, some evaluators may tend to follow paths through the system that lead to more scaffolding support than other paths, which can further confound the assessment of whether it is sufficient or not. Regardless of the possible interpretation issues, the low average points towards the possible need for improvement in this area. (Note: the implementation of changes in the questionnaire is discussed in Section 9.3 of Chapter Nine).

It should be noted that in both cases above, the lack of uniformity in ratings reflects on the wording used in the evaluation instrument and not necessarily on the prototype performance. This provides valuable feedback towards improving the evaluation instrument.

The highest average rating in this category was assigned to Evaluation Statement 2.1: 'The program provides assessment opportunities that are aligned with the objectives or outcomes'. The average rating of 4.7 follows from the assignment of four 5's and two 4's.

Spontaneous qualitative comments received from evaluators are:

- 'Colour coding of choices made could be clearer'. This refers to the colour coding used for the multiple-choice options and selected answers. Users select answers by clicking on buttons with text, presented as black text on blue buttons. Clicking on a button changes the colour of the button to indicate that it was selected as a possible answer. Clicking on it again deselects the button, which then changes back to the original colour. The colour used to indicate selection is a different shade of blue, but the evaluator suggests that the distinction would be clearer if a different colour was used.
- Evaluator D suggested that the wording of Evaluation Statement 2.2 'The assessment opportunities will serve to enhance trainees' performance' should be modified to read 'The assessment opportunities will serve to enhance trainees' ability to recognise hazards'. Although such a rephrasing would be more appropriate to the evaluation of the *ISGC* prototype, it would impact on the generic usability of the evaluation instrument by limiting it to hazard-specific systems.

8.5.1.2. *ISGC* heuristic evaluation results for Category 2

Table 8.25: Heuristic evaluation results of ISGC for Category 2: General Usability Heuristics.

CRITERIA	Evaluator rating						
	Usability Experts		Mining Experts		VR experts		
	A	B	C	D	E	F	Ave
1. Functionality							
1.1. The interface provides the level of functionality the user requires to complete a task.	5	3	5	4	4	4	**4.2**

1.2. The interface provides adequate back button functionality to return to a previous screen.	5	4	5	4	2	3	3.8
1.3. Icons, labels and symbols are intuitive and meaningful to trainees, bearing in mind the level of trainee context and experience.	5	4	5	4	2	3	3.8
Criterion average per evaluator	5	3.7	5	4	2.7	3.3	3.9
2. User guidance							
2.1. The interface provides clear indications of what the next required action will be.	3	4	5	4	3	4	3.8
2.2. Help for operating the program is accessible at any time and is appropriate.	2	4	5	2	2	2	2.8
2.3. Trainees receive clear instructions on how to use the training program.	5	4	5	4	3	3	4
2.4. Guidance to solve problems is given in the form of examples, diagrams, videos or photographs.	5	4	5	3	3	3	3.8
Criterion average per evaluator	3.8	4	5	3.3	2.8	3	3.6
3. Consistency							
3.1. There is consistency in the sequence of actions taken in similar situations.	5	5	4	4	4	5	4.5
3.2. There is consistency in the use of images, prompts, screens, menus, colours, fonts and layouts.	5	5	5	4	4	5	4.7
3.3. Objects, options, and permissible actions are visible so that users do not have to remember instructions.	5	5	5	5	4	5	4.8
3.4. Different screens that have similar operations, use similar elements for achieving similar tasks.	5	4	5	5	4	5	4.7
Criterion average per evaluator	5	4.8	4.8	4	4	5	4.7
4. Error Correction							
4.1. Error messages are expressed in plain language.	4	5	5	3	4	4	4.2
4.2. Learners are provided with the necessary help to recover from cognitive errors.	4	5	5	4	2	2	3.7
4.3. Error messages indicate precisely what the problem is and give simple, constructive, specific instructions for recovery.	3	5	5	4	2	2	3.5
Criterion average per evaluator	3.7	5	5	3.7	2.7	2.7	3.8
5. System Status							
5.1. The training program keeps the trainee informed about what is going on through constructive, appropriate and timely feedback.	5	4	5	4	3	3	4
5.2. For every action taken by the trainee, there is a visual or audio response by the training program so that learners can see and understand the results of their actions.	5	4	5	4	3	4	4.2
5.3. The program responds to actions initiated by the user and there are no surprise actions from the system's side.	5	4	4	5	4	4	4.3
Criterion average per evaluator	5	4	4.7	4.3	3.3	3.7	4.2

6. Aesthetics							
6.1. The screens are pleasing to look at.	5	5	5	5	3	4	4.5
6.2. The buttons and selections are of an adequately viewable size.	5	5	5	5	3	4	4.5
6.3. The text is of an adequately viewable size.	5	5	5	5	3	4	4.5
6.4. There is not too much content or information on the screens.	5	5	5	5	4	4	4.7
Criterion average per evaluator	5	5	5	5	3.3	4	4.6
7. Error Prevention							
7.1. The training program is designed in such a way that the learner cannot easily make serious errors.	5	5	4	4	3	5	4.3
7.2. When the learner makes an error, the system responds with an error message.	5	5	5	5	5	5	5
7.3. Trainees can recognise situations where errors are due to the way they provided input, and not due to incorrect content in their response.	5	5	5	5	4	4	4.7
7.4. The system is robust and reliable throughout.	5	5	4	4	4	5	4.5
Criterion average per evaluator	5	5	4.5	4.5	4	4.8	4.6
8. Interactivity							
8.1. The training program uses clear and simple terminologies that support trainees in understanding how to interact with the system.	4	4	5	5	4	4	4.3
8.2. The program provides interactions that support trainees in learning the necessary content.	5	4	5	5	3	4	4.3
8.3. Working through the program requires regular trainee interactivity to maintain attention and facilitate comprehension.	5	4	5	5	5	5	4.8
Criterion average per evaluator	4.7	4	5	5	4	4.3	4.5
Category Average per evaluator	4.6	4.5	4.9	4.3	3.4	3.9	
Overall average rating for Category 2							4.3

As indicated in Table 8.25, *ISGC* performed well in this category, with an overall average rating of 4.3 out of the maximum 5. Analysing the lower-rated aspects, four evaluation statements were rated on average as 3.8. Moreover, three evaluation statements received a lower rating:

- Evaluation Statement 2.2, 'Help for operating the program is accessible at any time and is appropriate', was rated 2.8. This rating includes four evaluator ratings of 2, which clearly indicate that the user assistance was deemed insufficient. Analysing this, however, reveals another problem: was the help provided inaccessible or was it inappropriate? Since the evaluation statement combined both accessibility and appropriateness in a single assessment, it is difficult to interpret this finding. An improvement to the evaluation instrument would be to separate accessibility and appropriateness into two evaluation statements. Regardless of the interpretation, the low rating warrants improvement in the user help provided by the prototype.

- The other two evaluation statements rated lower than 3.8 both belong to Criterion 4 (Error Correction):
  - Evaluation Statement 4.2, 'Learners are provided with the necessary help to recover from cognitive errors', was rated on average as 3.7.
  - Evaluation Statement 4.3, 'Error messages indicate precisely what the problem is and give simple, constructive, specific instructions for recovery', was rated with an average of 3.5.

  In both these cases it occurred again that the two VR evaluators assigned a 2 rating while two others assigned a 5, once again highlighting the possibility of some evaluators receiving more appropriate error-correction feedback due to taking paths through the prototype that were different from the paths taken by others. Nevertheless, the low averages show the need to revisit the prototype for possible improvement in this area.

Evaluation Statement 7.2, 'When the learner makes an error, the system responds with an error message', is the only evaluation statement to receive a perfect 5 average rating in the complete *ISGC* evaluation.

Two further comments were received for this category:

- Evaluator A commented that the size of text in error messages was rather small and should be increased.
- Evaluator C assessed it as 'Excellent!'

8.5.1.3. *ISGC* heuristic evaluation results for Category 3

Table 8.26: Heuristic evaluation results of ISGC for Category 3: VR System Design Heuristics.

CRITERIA	Evaluator rating						
	Usability experts		Mining experts		VR experts		
	A	B	C	D	E	F	Ave
1. User control							
1.1. The user is able to interact with, or control, the virtual environment in a natural manner.	3	4	2	2	2	2	**2.5**
1.2. Responses from the environment to the participant's control actions and movements, are perceived as immediate or close-to-immediate.	4	4	2	4	2	4	**3.3**

1.3. The system permits easy reversal of actions.	4	4	2	2	2	3	2.8
1.4. Trainees are able to exit the system at any time when they need to do so.	4	4	1	5	2	1	2.8
Criterion average per evaluator	3.8	4	1.8	3.3	2	2.5	2.8
2. Multimodal System output / feedback							
2.1. The effect of the trainee's actions on objects in the virtual environment, is immediately visible and conforms to the laws of physics and the trainee's perceptual expectations.	4	5	5	4	2	3	3.8
2.2. The visual representation of the virtual world maps to the trainee's normal perception of that environment.	5	5	5	4	4	3	4.3
2.3. Distortions are not noticeable in visual images.	5	5	4	5	4	3	4.3
2.4. Audio is integrated seamlessly into user task activity.	5	5	5	5	3	3	4.3
2.5. Audio information is meaningful and timely.	5	5	5	5	4	4	4.7
Criterion average per evaluator	4.8	5	4.8	4.6	4.2	3.2	4.3
3. Presence							
3.1. Users feel as if they are part of the virtual environment and not isolated from it.	4	5	5	5	3	4	4.3
3.2. The virtual environment experience is consistent with similar real-world experiences.	4	5	5	5	3	4	4.3
Criterion average per evaluator	4	5	5	5	3	4	4.3
4. Orientation							
4.1. Users do not find it difficult to maintain knowledge (or 'awareness') of their location while moving through the virtual environment.	5	5	4	5	4	4	4.5
4.2. The virtual environment includes appropriate spatial labels and landmarks to assist user orientation.	5	5	5	4	2	5	4.3
4.3. It is clear to the user how to exit the virtual environment.	5	5	5	4	2	2	3.8
Criterion average per evaluator	5	5	4.7	4.3	2.7	3.7	4.2
5. Navigation							
5.1. Is it easy for users to move and reposition themselves in the virtual environment.	4	5	2	4	2	3	3.3
5.2. Ways of navigating are consistent throughout the system.	5	5	5	3	4	5	4.5
Criterion average per evaluator	4.5	5	3.5	3.5	3	4	3.9
6. Object interaction: selection and manipulation							
6.1. Input devices are easy to use and easy to control.	5	5	5	4	3	3	4.2
6.2. Object interactions are designed realistically to reproduce real-world interaction.	5	5	4	4	3	3	4.0
6.3. The system provides the ability to rotate 3D objects and increase detail levels when necessary for task performance.	3	3	1	1	1	1	1.7
Criterion average per evaluator	4.3	4.3	3.3	3	2.3	2.3	2.5

7. Fidelity							
7.1. The simulations in the system are accurate.	5	5	5	5	2	3	**4.2**
7.2. The objects in the virtual environment move in a natural manner.	5	5	5	5	2	3	**4.2**
7.3. The virtual environment displays adequate levels of realism.	5	5	5	4	3	3	**4.2**
7.4. High-fidelity graphics are used where required.	4	5	3	5	2	3	**3.7**
Criterion average per evaluator	4.8	5	4.5	4.8	2.3	3	**4.1**
8. Various user modes							
8.1. The system employs various user modes to cater for a range of users from novices to experts.	3	3	3	1	1	2	**2.2**
8.2. The system provides various user-guidance modes, e.g. Free mode, Presentation mode, Guided mode and Discovery mode.	2	3	1	1	1	2	**1.7**
Criterion average per evaluator	2.5	3	2	1	1	2	**2**
Category Average per evaluator	4.3	4.6	3.8	3.8	2.5	2.0	
Overall average rating for Category 3							**3.7**

Table 8.26 contains the results of the heuristic evaluation of the *ISGC* prototype for Category 3 – VR System Design Heuristics. Of the four categories evaluated, this category scored the lowest average, 3.7, compared to 4.1 for Category 1 and 4.3 for both Categories 2 and 4. Specifically, the average ratings per evaluator are very low for the two VR experts, 2.5 and 2, respectively.

There are several evaluation statements that were rated, on average, less than 3. These include three statements related to user control and both statements relating to user modes. This highlights certain aspects where the *ISGC* prototype can be improved.

- 1.1: 'The user is able to interact with, or control, the virtual environment in a natural manner'. The low rating here, namely 2.5, indicates that there is a lack of user control. It must be explained that, due to the inadequate computer-based training experience of the envisaged training audience, a design decision was taken to explicitly limit the user control. The intention was to prevent situations where users might become more concerned about how to navigate or how to control the objects in the virtual environment than about focusing on the learning. The ability to control the virtual environment, however, is an important feature in VR systems and the prototype should be modified to offer more user control as trainees become more confident in using the computer as a training medium.
- 1.3: 'The system permits easy reversal of actions'. An average rating of 2.8 indicates clearly that this aspect is inadequate in the prototype.

- 1.4: 'Trainees are able to exit the system at any time when they need to do so'. The low rating here of 2.8 indicates that there is not always an exit option available. Once again, it was an explicit design decision not to permit users to exit the program at all stages. For example, a constraint prevents users from exiting immediately after identifying a hazard. They have to first either correct the hazard or view the possible consequence. This ensures that users deal with the hazard without disruption. Nevertheless, an exit option could be available between hazards.
- 6.3: 'The system provides the ability to rotate 3D objects and increase detail levels when necessary for task performance'. This evaluation statement received a very low average rating of 1.7. However, in this particular prototype, dealing with geological features, there is no need to physically rotate objects as the user can see only the hanging wall or sidewall and is not allowed a view 'inside' a rock. This ensures realistic simulation of the underground conditions, where a geological condition must be identified based on what the trainee can see on the outside of the rock. This statement was included in the evaluation framework due to its generic applicability to other VR systems.
- 8.1: 'The system employs various user modes to cater for a range of users from novices to experts'. The average rating for this evaluation statement was 2.2. The omission of this aspect has already been highlighted in Evaluation Statement 5.2 of Category 1: 'The system caters for trainees with different levels of expertise regarding the content'. To avoid duplication in the evaluation framework, these two evaluation statements should be combined into one.
- 8.2: 'The system provides various user-guidance modes, e.g. Free mode, Presentation mode, Guided mode and Discovery mode'. The statement was consistently rated low, with an average rating of 1.7. Due to the limitations placed on user control, *ISGC* did not make provision for a free or discovery mode, whereby users could freely move around in the virtual environment and manipulate all the objects.

Even though some of the above aspects were deliberately restricted in the prototype due to the perceived lack of computer usage ability among trainees, it remains important that an evaluation instrument for VR systems should include these aspects, since they are powerful features that would enhance the learning experience.

The highest average rating for this category, 4.7, was allocated to Evaluation Statement 2.5: 'Audio information is meaningful and timely'.

There were no open-ended responses received from the evaluators for this category.

8.5.1.4. ISGC heuristic evaluation results for Category 4

Table 8.27: Heuristic evaluation results of ISGC for Category 4: Context-specific Heuristics.

CRITERIA	Evaluator rating						
	Usability experts		Mining experts		VR experts		
	A	B	C	D	E	F	Ave
1. Authentic tasks							
1.1. The training system supports particular work practices in the context of their natural work environment.	5	5	5	5	3	4	4.5
1.2. The system is customised according to learner-specific needs and the relevance of the curriculum.	5	5	5	4	3	4	4.3
1.3. The program includes tasks applicable to the actual job context of the trainee.	5	5	3	4	3	4	4
Criterion average per evaluator	5	5	4.3	4.3	3	4	4.3
2. Appropriate reference materials							
2.1. The system includes additional reference materials, providing information to trainees on standard operating procedures used in the application domain.	4	3	1	4	1	3	2.7
2.2. The reference materials included in the system are relevant to the problem scenarios.	4	4	1	4	1	3	2.8
2.3. The reference materials are at a level appropriate to the trainees.	4	4	1	4	1	4	3
Criterion average per evaluator	4	3.7	1	4	1	3.3	2.8
3. Comprehensive scope							
3.1. The learning material in the program covers all the vital aspects relating to the topics being addressed.	4	5	4	5	3	4	4.2
3.2. The training also covers possible consequences of trainees not applying the learning material correctly in their work place.	5	5	3	5	4	5	4.5
Criterion average per evaluator	4.5	5	3.5	5	3.5	4.5	4.4
4. Adaptive design							
4.1. The design of the training system is adaptive to changes in site practices.	5	5	5	5	4	5	4.8
4.2. The system refers to the latest current standard operating procedures.	5	5	5	5	4	5	4.8
4.3. The system randomises assessment details such as questions and multiple choice answers when presenting assessment opportunities to trainees.	5	3	5	5	4	4	4.3
Criterion average per evaluator	5	4.3	5	5	4	4.7	4.6

5. Appropriate record keeping							
5.1. The system maintains student records and assessment results.	5	5	5	3	5	5	4.7
5.2. The system monitors and displays student progress.	5	5	5	2	3	4	4
5.3. The system ensures legal compliance in the application domain by capturing detailed individual performance data.	5	5	5	3	5	5	4.7
Criterion average per evaluator	5	5	5	2.7	4.3	4.7	4.5
6. Trainee preparedness							
6.1. Trainees are shown how to use the software prior to doing the training program.	5	5	5	5	4	4	4.7
6.2. PC literacy pre-training is available to trainees not comfortable with using computers for training.	5	3	5	5	5	5	4.7
Criterion average per evaluator	5	4	5	5	4.5	4.5	4.7
7. Relevant subject matter							
7.1. The subject matter matches the goals and objectives of the training program.	5	5	5	4	4	5	4.7
7.2. The subject matter is presented in an appropriate content structure.	5	5	5	5	3	5	4.7
7.3. The information provided in the program is accurate.	5	5	4	5	4	5	4.7
7.4. The system 'speaks the trainee's language' by using terms, phrases, symbols and concepts familiar to the trainee and common to the application domain.	5	5	4	5	2	5	4.3
7.5. The level of language use, in terms of grammar and style, is applicable to the target audience.	5	5	4	5	3	5	4.5
Criterion average per evaluator	5	5	4.4	4.8	3.2	5	4.6
8. Understandable and meaningful symbolic representation							
8.1. Symbols, icons and terminology used to represent concepts and objects are used consistently throughout the program.	5	5	5	4	4	4	4.5
8.2. Symbols, icons and terminology used are intuitive within the context of the task.	5	5	5	4	4	3	4.3
8.3. Metaphors used correspond to real world objects or concepts.	5	5	5	4	4	3	4.3
Criterion average per evaluator	5	5	5	4	4	3.3	4.4
Category Average per Evaluator	4.8	4.7	4.2	4.3	3.4	4.3	
Overall average rating for Category 4							4.3

As shown in Table 8.27, the overall average rating for this category is 4.3, which together with Category 2, is jointly the highest of the four categories. Within the category, the highest average rating of 4.8 was assigned to two evaluation statements in Criterion 4 (Adaptive design), namely 4.1 'The design of the training system is adaptive to changes in site practices' and 4.2 'The system refers to the latest current standard operating procedures'.

The only notable 'poor performers', scoring less than 4 were all three evaluation statements of Criterion 2 (Appropriate reference materials). The average for this criterion is only 2.8, indicating a lack in the prototype in terms of reference materials. Most of the open-ended responses received for this category were linked to this aspect:

- Evaluator A: 'There are no additional reference materials'.
- Evaluator C: 'Additional photos should be made available on each hazard'.
- Evaluator D: 'The system should link to the relevant SOPs'.

It would be beneficial to the learners to be able to access the mine's standard operating procedures (SOPs), as these documents contain the latest standards applicable to dealing with and supporting the various geological conditions. This poses an additional challenge, however, as these documents are regularly updated and the system would have to link to the latest version of the document so as to present current information. This implies that a procedure should be put in place to ensure that the system points to the latest version of an updated SOP.

Including additional photographs of similar geological conditions would assist not only in identifying a hazard, but would also expose trainees to further examples of such a hazard, and would therefore strengthen the system.

The only other qualitative comment received was from Evaluator D, indicating that, although the record keeping in the prototype (Criterion 5) was appropriate, the assessment results are not categorised and trainers may have difficulty in identifying problem areas. This comment refers to the evaluation feedback report provided by *ISGC,* in which the trainee performance is indicated per hazard and the trainee's final score is given. An improvement and extension to this report would be information on trainee performance in terms of hazard identification, associated risks and control measures, and appropriate support for each hazard. Such information would allow the trainer to determine whether a trainee struggles with identification, controls or support in general, over and above his/her performance related to a particular geological condition.

8.5.1.5. ISGC heuristic evaluation overall results

Table 8.28 presents the average rating allocated by each evaluator, calculated as the total scores allocated by each evaluator divided by the total number of evaluation statements (105). The overall average of all six evaluators for the complete evaluation is 4.1 out of 5, which can be viewed as 82%, 6% higher than the ratings of the *LSF* prototype, shown in Table 8.7.

Table 8.28: Overall averages per evaluator for the ISGC prototype

Evaluator	A	B	C	D	E	F	All
Average	4.6	4.5	4.5	4.1	3.1	3.8	4.1

Table 8.29 summarises the overall average ratings of *ISGC* per category. Categories 2 (General Usability) and 4 (Context-specific heuristics) were both allocated the highest average of 4.3 (86%) while, similarly as was reported for *LSF* in Table 8.8, Category 3 (VR System Design) has the lowest average of 3.7 (74%). The category averages are discussed in more detail in Section 8.6.1.

Table 8.29: Overall averages per category for the ISGC prototype.

Category	1	2	3	4	All
Average	4.1	4.3	3.7	4.3	4.1

8.5.2. Findings of the user satisfaction survey of the ISGC prototype

The user satisfaction questionnaire was completed by 52 participants. Since the same instrument was used as for the *LSF* evaluation, it was decided to focus this evaluation on higher ranks of underground workers, that is, miners, shift supervisors and mine overseers. The next three sections discuss the findings of the survey with regard to the biographical details of the participants, information related to the *ISGC* prototype, and the open-ended section on user comments, respectively. Interpretation of the findings is given where the findings are discussed.

8.5.2.1. Biographical information

Table 8.30 lists the three job titles and the number of participants holding each role. The rank of miner is reached after specific mining training and successful performance evaluation. To qualify for miner training, a mine worker must have experience as a panel operator, winch driver and rock drill operator. After miner, the next level of promotion is shift supervisor (previously known as shift boss), and then mine overseer (previously mine captain).

Table 8.30: ISGC participant job titles, number and percentage of incumbents.

Job title	Number	Percentage
Miner	45	86.5
Shift supervisor	4	7.7
Mine overseer	3	5.8
Total	52	100

Table 8.31 indicates the school education level completed by the participants, showing that 57.7% of them had completed secondary education (Grade 12 level). This is in line with the data in Table 8.10 from the *LSF* survey.

Table 8.31: Number and percentage of ISGC participants' schooling levels.

School level completed	Number	Percentage
< Grade 6	1	1.9
Grade 6 – 7	3	5.8
Grade 8 – 9	5	9.6
Grade 10 – 11	10	19.2
Grade 12	30	57.7
No response	3	5.8
Total	**52**	**100**

In terms of exposure to technology, Table 8.32 indicates how many participants had previously used computers, cell phones and bank ATMs. Thirty-seven of the 52 participants (71.2%) had prior exposure to the use of computers before doing the *ISGC* training, which is considerably higher than the 34% indicated by the *LSF* participants. The *ISGC* participants' exposure to cell phones and bank ATMs was still notably higher than their use of computers. Table 8.33 indicates that 32 of them (61.5%) had done the pre-training on using the mouse prior to doing the training prototype.

Table 8.32: ISGC participant exposure to technological devices.

Device used before	YES	YES %	NO	NO %
Computer	37	71.2	15	28.8
Cell phone	49	94.2	3	5.8
Bank ATM	47	90.4	5	8.6

Table 8.33: ISGC participants choosing to do pre-training.

Pre-training done?	Number	Percentage
YES	32	61.5
NO	20	39.5
Total	52	100

8.5.2.2. Data regarding the *ISGC* prototype

Table 8.34 indicates the home language of the participants as well as the language they selected at the start of the *ISGC* training, with regard to the four languages for which the *ISGC* prototype makes provision.

From the information presented in Table 8.34 it is evident that:

- The four languages available in the *ISGC* prototype apply only to the home language of 26 of the participants (50%), with the other 26 (50%) having a home language not available in the prototype. These participants had to choose one of the four available languages, which explains why more participants chose English for the *ISGC* prototype. The other home languages indicated, were Afrikaans (14), isiZulu (3), Tsonga (7) and South Sotho (2).
- Some Tswana, Sepedi and Xhosa speakers actually chose to do the prototype in English, even though Tswana, Sepedi and Xhosa were available to them. Once again, this is probably related to the use of English as the main language of instruction at the mine. Trainees might have felt more comfortable using a training system containing mining terminologies in English.

Table 8.34: Participants' home languages and selected ISGC language.

Home Language	Number	%	*ISGC* selected language	Number	%
English	5	9.6	**English**	45	86.5
Xhosa	3	5.8	**Xhosa**	2	3.8
Tswana	12	23.1	**Tswana**	5	9.6
Sepedi	6	11.5	**Sepedi**	0	0
Other	26	50	-	-	-
TOTALS	**52**	**100**		**52**	**100**

As was done in Section 8.4.2.2 for the *LSF* prototype, the findings related to the *ISGC* prototype (Questions 2.2 to 2.20 of the questionnaire), were processed under various usability-related categories. The question numbers from the questionnaire relating to each category are as follows:

- Ease of use: 2.4, 2.5, 2.6, 2.7, 2.10 and 2.14;
- Learnability: 2.9 and 2.15;
- Satisfaction: 2.2, 2.3 and 2.16;
- Authenticity: 2.8, 2.11, 2.12, 2.13 and 2.17; and
- Method of choice: 2.18, 2.19 and 2.20.

The seven tables in the subsections following present the findings per category. Ratings are indicated in percentages. All 52 participants responded to all the questions, except for Questions 2.5 and 2.16 where only 51 responses were received. In the categories where these two questions belong, an additional column is provided, indicating NR for No Response. An average column (AVE) shows the average Likert scale rating for each question and the average for the category. As mentioned in Section 8.5.2.2, please note that in the rating system for the user satisfaction survey, 1 is the most positive option and 5 the most negative, which is the opposite of the rating system used for the heuristic evaluation. Consequently, where a rating of 4 to 5 would have been the preferred response in the heuristic evaluation, a rating of 1 to 2 would be positive feedback for the user satisfaction survey.

Ease of use

The responses received for Questions 2.4, 2.5, 2.6, 2.7, 2.10 and 2.14 are grouped together in an ease-of-use category. Table 8.35 summarises the findings for the questions related to the ease of use of *ISGC*. As was the case with the findings of the *Ease of use* category for the *LSF* prototype in Table 8.14, the results for Question 2.6 are presented in reverse order in Table 8.35. This way, the most positive responses are indicated first, which is in line with the order of all the other responses, and uniform data is provided for statistical analysis.

As was the case with the *LSF* evaluation, Table 8.35 shows that most responses received were in the first and second scale options, representing *Very much/Very easy/etc.* and *much/easy/etc.* respectively.

As indicated in Section 8.5.2.1, 71.2% of the participants in the *ISGC* evaluation had previously used computers, compared to the 33.8% in the case of the *LSF* evaluation. Even so, most participants still required some form of assistance from the facilitator. If facilitator assistance (Question 2.6) and time given for completion (Question 2.14) are omitted from the evaluation of ease of use, the total percentages in the first and second scale options accumulatively for Questions 2.4, 2.5, 2.7 and 2.10 are 50%, 78.8%, 86.6% and 71.1% respectively, which represents an average of 71.5% positive responses for the 52 users. The low accumulative score of 50% for Question 2.4, compared to the others in this category, is a concern.

Table 8.35: Responses for the Ease of use category – ISGC prototype.

	Rating:	1	2	3	4	5	NR	AVE
2.4.	**How easy was this training program to use?**	17.3	32.7	32.7	15.4	1.9	0	2.5
2.5.	**How easy was it to work with the mouse?**	40.3	38.5	13.5	5.8	0	1.9	1.8
2.6.	**How much assistance did you require from the facilitator?**	25	30.8	25	11.5	7.7	0	2.5
2.7.	**How well could you understand the questions in the program?**	30.8	55.8	9.6	3.8	0	0	1.9
2.10.	**How much are you at ease using computers for training?**	42.3	28.8	17.3	11.5	0	0	2.0
2.14.	**Were you given enough time to complete the training program?**	44.2	30.8	15.4	1.9	7.7	0	2.0
	Average:	**33.3**	**36.2**	**18.9**	**8.3**	**2.9**	**0.3**	**2.1**

Learnability

Table 8.36 presents the responses relating to the learning value of the system, with 84.6% answering that they have learnt *Very much* or *Much* (Question 2.9). For Question 2.15, all of the participants (100%) indicated that the program had made them *Very Much* or *Much* more aware of hazards in the workplace.

Table 8.36: Responses for the Learnability category – ISGC prototype.

	Rating:	1	2	3	4	5	AVE
2.9.	**How much did you learn by using this program?**	44.2	40.4	9.6	5.8	0	1.8
2.15.	**Will this training program help you to be more aware of the hazards in the workplace?**	61.5	38.5	0	0	0	1.4
	Average:	**52.9**	**39.5**	**4.8**	**2.9**	**0**	**1.6**

Satisfaction

Table 8.37 shows answers related to enjoyment, interest and feedback satisfaction, which are combined in the *Satisfaction* category.

Table 8.37: Responses for the Satisfaction category – ISGC prototype.

Rating:	1	2	3	4	5	NR	AVE
2.2. How interesting was this training program to you?	50	36.5	11.5	0	1.9	0	2.2
2.3. How much did you enjoy doing this program on the computer?	57.7	26.9	15.4	0	0	0	1.6
2.16. How satisfied are you with the feedback that you received from the program while you were doing the training?	40.4	50	5.8	1.9	0	1.9	1.7
Average:	**49.4**	**37.8**	**10.9**	**0.6**	**0.6**	**0.6**	**1.8**

Most responses received were in the first two scale options. For Question 2.2, the 50% (26) responses for option 1 plus the 36.5% (19) for option 2, indicating *Very interesting* and *Interesting* respectively, represent 86.5% of the participants. For Question 2.3, 84.6% (44) of the participants indicated their enjoyment as *Very much* or *Much*, while 90.4% (47) indicated in response to Question 2.16 that they were *Very satisfied* or *Satisfied* with the system's feedback. From a negative perspective, only 1.9% (1 participant) found the program *Not at all interesting* and, similarly, only 1.9% (1 participant) was *Not really satisfied* with the system feedback.

Authenticity

Five questions are grouped together in the *Authenticity* category, namely Questions 2.8, 2.11, 2.12, 2.13 and 2.17. The responses to Questions 2.8, 2.13 and 2.17 are summarised in Table 8.38, and the responses to Questions 2.11 and 2.12 in Table 8.39. As was explained in the discussion of the *LSF* prototype, the reason for separating Questions 2.11 and 2.12 in the analysis, was the way they were asked in the questionnaire. For these two questions, a four-option scale was used, because the nature of the questions did not make an *Average* option feasible. Questions 2.11 and 2.12 dealt with the occurrence of the accidents shown in the system and the options were *None of them*, *Some of them, Most of them* and *All of them*.

Table 8.38: Responses to Questions 2.8, 2.13 and 2.17 of the Authenticity category – ISGC prototype.

	Rating:	1	2	3	4	5	AVE
2.8.	**How easily did you recognise the objects on the screen?**	21.2	44.2	30.8	3.8	0	2.2
2.13.	**How realistic were the accidents that you saw in this training program?**	26.9	51.9	17.3	3.8	0	2.0
2.17.	**To what extent are the geological hazards shown in this program relevant to your job?**	51.9	28.8	17.3	1.9	0	1.7
	Average:	**33.3**	**41.6**	**21.8**	**3.2**	**0**	**2.0**

Once again, the responses in Table 8.38 follow a similar trend to the previous categories, in that most of the responses were option 1 or 2. In terms of the recognisability of the graphic objects in *ISGC* (Question 2.8), 65.4% of the responses were option 1 or 2. With a more positive response, 78.8% of the participants selected option 1 or 2 to assess the realism of the accidents (Question 2.13), and 80.7% indicated likewise for the relevancy of the hazards (Question 2.17).

Table 8.39: Responses to Questions 2.11 and 2.12 of the Authenticity category – ISGC prototype.

	Options:	None	Some	Most	All
2.11.	**Do you believe that the accidents you saw in the program can really happen?**	9.6	23	30.8	36.5
2.12.	**Do you believe that the accidents you saw in the program can really happen to you?**	26.9	42.3	15.4	15.4
	Average:	**18.3**	**32.7**	**23.1**	**26**

As was the case with *LSF*, the responses to Questions 2.11 and 2.12, shown in Table 8.39, indicate that even though 67.3% of the participants agreed that *Most* (30.8%) or *All* (36.5%) of the accidents shown could really happen, only 30.8% believed that *Most* (15.4%) or *All* (15.4%) of these accidents could actually happen to them. Of particular concern is that 26.9% of the participants believed that *None* of the accidents could happen to them. This is alarming, since the geographical conditions selected for the system were based on a case study of previous incidents at that particular mine, as explained in Section 7.4.1.

Method of choice

The responses to this category on the computer as training medium, are given in Table 8.40. All 52 participants responded to both questions. In response to Question 2.18, 86.5% of participants indicated that they would *Very much* (61.5%) or *Much* (25%) like to do similar training on the computer again. None responded that they did not want to do similar computer-based training again. Similarly, in response to Question 2.19, 88.5% selected option 1 or 2 to indicate their preference for this type of training, rather than classroom training only.

Table 8.40: Responses to Questions 2.18 and 2.19 for the Method of choice category – ISGC prototype.

	Rating:	**1**	**2**	**3**	**4**	**5**	**AVE**
2.18.	**How much would you like to do training like this on the computer again?**	61.5	25	11.5	1.9	0	1.5
2.19.	**Do you think this type of training on the computer is better than just listening to an instructor in the classroom?**	57.7	30.8	5.8	1.9	3.8	1.6
	Average:	**59.6**	**27.9**	**8.7**	**1.9**	**1.9**	**1.6**

Table 8.41 lists the participants' preferred method of training. Only two participants (3.8%) preferred classroom lectures. Similarly, only two participants selected video-based training. Many selected computer-based training (44.2%) or the combination of lectures and computer-based training (25%) as their preferred training option. As explained in the discussion of this category in the *LSF* findings (Section 8.4.2.2), the researcher does not advocate replacing the essential practical, hands-on training by the other specified modes, but rather that the other training modes should be used to supplement practical training.

Table 8.41: Responses to Question 2.20 for the Method of choice category – ISGC prototype.

	Options:	**Classroom lecture**	**Practical**	**Video**	**Computer**	**Lecture and Computer**
2.20.	**Please indicate your preferred method of training**	3.8	23.1	3.8	44.2	25

In summary, similarly to the user responses received for the *LSF* prototype, it is evident from Tables 8.35 to 8.41 that the trainees rated the *ISGC* training prototype highly. The next section deals with the qualitative comments received via open-ended questions.

8.5.2.3. Qualitative user comments on the *ISGC* prototype

The third part of the user satisfaction questionnaire contained five open-ended questions. To assist the reader, the open-ended questions are repeated:

1. What do you think are the best features of the program?
2. What aspects of the program do you think should be improved?
3. Is there any other training that you would prefer to do on the computer?
4. Please describe problems you encountered in using the *Interactive Simulated Geological Conditions* program.
5. Do you have any other comments on the *Interactive Simulated Geological Conditions* program?

Question 1 received varied responses, ranging from the advantages of visualising the potential effects of incorrectly addressing geological hazards through to praising the sound effects. The feature highlighted by most participants was the relevance of the content and the associated risks within their workplace. Some specific comments were:

- 'It refreshes my brain on geological standards and procedures'.
- 'The program is easily understandable'.
- 'The graphics are very realistic'.
- 'The best feature of the program is it gives correct answers if you made mistake'.

For Question 2, only four suggestions were received for possible improvements to the program. They related to four different issues:

- 'Ask the student, are you sure about the answer before pressing Next button'. Out of both groups of participants that completed the user satisfaction questionnaires (*LSF* and *ISGC*), only one participant made this valid request.
- 'Visibility of some incidents can be improved'. The participant did not indicate which scenes should be improved, hence it was vague.
- 'Some answers I don't agree with'. The participant did not indicate the answers with which he disagreed, and was the only individual who gave such feedback. All the answers had been thoroughly checked by training staff prior to implementation, so one can only assume that this participant had not mastered the subject matter.

- An interesting suggestion received, was 'Maybe a little read through of standards before doing the program would help, as some people are exposed to the standards of this mine for the very first time when doing this program'. This should certainly not be the case, but the biographic details of this participant revealed that he was new to that particular mine, although he had 32 years of experience and functions as a mine overseer.

Question 3 dealt with the identification of other training that could be delivered via computers. As was the case with *LSF*, many participants answered 'anything' or 'all training on mining practices'. This indicates a preference for computer-based training, as was also shown in the responses to Question 2.20 – see Section 8.5.2.2. Other, more specific, answers include: mine overseer ticket training, rock engineering practicals, all ex-leave training and teamwork.

In Question 4 the participants were asked to mention problems they had encountered while using *ISGC*. Although there were 52 participants, only three comments were received. Two referred to difficulty in understanding the English language. Biographic details revealed that one was a Tswana speaker and the other a Tsonga-speaking miner who had both done the program in English. The third comment received was 'This is giving me a lot of stress', without elaboration. The stress might refer to the participant not knowing the subject matter well enough to answer the questions with confidence or it could refer to the high realism of the fall-of-ground visuals that had a shock effect on some of the participants.

Question 5 required participants to write down any other comments about the system. Only positive comments were received, for example,

- 'Great program',
- 'Good program',
- 'Keep up the good work!!!',
- 'It was good for me', and
- 'Very well set up'.

8.5.2.4. Discussion and interpretation of findings of *ISGC*

The responses to the usability-related categories Ease of use, Learnability, Satisfaction and Authenticity were presented in Tables 8.35, 8.36, 8.37 and 8.38, respectively. As indicated in these tables, the overall average ratings for these categories were:

- Ease of use: 2.1;

- Learnability: 1.6;
- Satisfaction: 1.8; and
- Authenticity: 2.0.

According to the rating system used, a 1 rating is the most positive rating. It is therefore clear that, as was the case during the *LSF* evaluation, the participants rated Learnability as the best usability trait, slightly better than Satisfaction, which is followed by Authenticity and Ease of use. Unlike *LSF*, these four averages are closer to a 2 rating than to a 1 rating, meaning that all of these usability traits were evaluated more as positive than very positive.

The high ratings for Learnability are encouraging, particularly the fact that all of the participants responded with a 1 or a 2 rating to indicate that the program had made them either *Very Much* or *Much* more aware of hazards in the workplace. The slightly lower rating for Authenticity is a concern, specifically due to the fact that one of the main focus areas in the development of *ISGC* was to improve the realism of the graphics. The average rating received for Authenticity of 2.0 thus indicates that there is still room for improvement in this area.

For the open-ended section, far fewer comments were received than in the *LSF* user satisfaction survey. This is due to the fact that there were only 52 participants, in comparison to the 195 of *LSF*. Moreover, many of the 52 participants did not submit any answers to the open-ended questions. As indicated in Section 8.5.2.3, the responses received did not clearly identify any areas for improvement, but most of the comments indicated high praise for the system.

8.5.2.5. Internal consistency reliability: *ISGC* questionnaire responses

As was done with the responses received for the *LSF* questionnaire in Section 8.4.2.5, the internal consistency reliability of the user satisfaction questionnaire was also verified for the *ISGC* responses. Similarly, because the number of items should not be less than four when verifying internal consistency reliability and calculating perception measures, the only usability category in the questionnaire with sufficient items that could be evaluated for internal consistency reliability is the *Ease of use* category, using responses to the following questions:

2.4: How easy was this training program to use?
2.5: How easy was it to work with the mouse?
2.7: How well could you understand the questions in the program?

2.10: How much are you at ease using computers for training?

Internal consistency reliability

The results of the scale reliability test as conducted on the subset of the above four questionnaire item responses, yielded a 'good' Cronbach alpha value of 0.76. Thus Questions 2.4, 2.5, 2.7 and 2.10 show a combined internal consistency reliability. The mean responses of these response ratings for each participant can therefore be used as a measure of participants' perception of ease of use.

The calculation of the ease-of-use score

Table 8.42 reports the ease-of-use mean scores calculated for the categories of *schooling level* (responses for Question 1.1, indicated in column 'schooling') and *pre-training* (responses for Question 1.3, indicated in column 'pre-training'), as well as the standard deviation for each grouping.

Table 8.42: Means table for the ease-of-use analysis for ISGC responses.

Schooling	Pre-training	N	Mean	Standard Deviation	Minimum	Maximum
1	1	1	3.00	.	3.00	3.00
2	0	1	3.50	.	3.50	3.50
	1	2	2.50	0	2.50	2.50
3	0	3	2.00	0.25	1.75	2.25
	1	2	2.13	0.53	1.75	2.50
4	0	6	2.21	0.91	1.00	3.25
	1	4	2.13	0.75	1.25	2.75
5	0	9	2.03	0.69	1.00	3.25
	1	21	1.71	0.51	1.00	2.75

In the 'schooling' column, values of 1 to 5 are used to indicate the five schooling levels in the questionnaire, as was indicated in Table 8.31. In the pre-training column, 0 indicates that pre-training was not done and 1 indicates that pre-training was done. N indicates the number of responses received for each interaction of *schooling* and *pre-training*.

Table 8.42 shows that the largest group of participants (21) did have Grade 12, but did not do the pre-training. This group also has the lowest mean score (1.71), which indicates that they found the system between 'very easy' and 'easy' to use.

General linear model approach to analysis of variance

Table 8.43 presents the results of an analysis of variance on the ease-of-use perception scores, where the effects of schooling level and pre-training are included in the analysis of variance model to evaluate the statistical significance of these effects on perceptions of how easy it is to use the computer training system.

As indicated in Table 8.42, some categories and combinations of categories (schooling by pre-training) have very few responses. As a reliable analysis of variance cannot be conducted on scores with such sparsely populated cells (low frequencies), some categories were combined to be able to calculate analysis of variance reliably. For the schooling level variable, all categories less than Grade 12 were combined into one category, leaving just two categories for the analysis: schooling level < Grade 12 and schooling level = Grade 12.

Table 8.43 has 3 separate blocks:

- Block 1 indicates the levels of the effects included in the model, representing the independent variables, 'schooling' (with values 1 and 2) and 'pre-training' (with values 0 and 1).
- Block 2 indicates that the responses of 49 participants were used due to some missing responses.
- Block 3 reports on the analysis of variance itself, indicating that the general F statistic for the test of 2.8 is significant on the 5% level of significance.

Table 8.43: GLM analysis of variance for ease-of-use for ISGC responses.

Class	Levels	Values
Schooling	2	1 2
Pre-training	2	0 1

Number of Observations Read	52
Number of Observations Used	49

Source	Degrees of Freedom	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	3.316	1.105	2.80	0.0505
Schooling	1	1.901	1.901	4.82	0.0333
Pre-training	1	0.216	0.216	0.55	0.4626
Schooling*Pre-training	1	0.320	0.320	0.81	0.3724
Error	45	17.745	0.394		
Corrected Total	48	21.061			

As indicated in Block 3 of Table 8.43, the effect of schooling on ease-of-use perceptions is statistically significant on the 5% level of significance. Table 8.44 indicates the nature of the effect of schooling level on ease-of-use perceptions. Since only two training perception level means are involved in this instance, the deduction can be made that schooling level does ease the use of the *ISGC* prototype. This follows from comparing the perception means of 2.29 and 1.8, the latter being a better ease-of-use rating according to the Likert scale used, indicating that the higher a participant's schooling level, the easier it was to use the system.

Table 8.44: Ease-of-use perception means for the two schooling levels

Significance indicator	Mean	N	Schooling level
A	2.29	19	< Grade 12
B	1.80	30	Grade 12

The results of the statistical analysis in this section indicated that higher schooling levels contribute to the ease with which *ISGC* is used, regardless of pre-training. As was discussed in Section 8.4.2.5, the schooling level of entrants into the mining industry is increasing drastically, but there is still a large cohort of existing employees with low schooling levels.

Regardless of this finding, the overall evaluation of *ISGC*, including ease of use, was still very positive.

8.6. Comparison of findings

This section compares the results of evaluating the *LSF* prototype to the *ISGC* results. The value of such a comparison has limitations, because the *ISGC* prototype was developed not only as an improvement of the *LSF* prototype, but also as an exploratory training system utilising alternative development tools to improve realism and incorporating a shift in focus from generic hazards to specific hazards (geological conditions). *ISGC* did, however, incorporate usability improvements identified during designer reflection in DBR Cycles 1 and 2. The results of the evaluation and additional feedback received, as described in this chapter, emphasise the value of the evaluation framework in providing an instrument with which to evaluate such systems in terms of the specified categories, by providing quantitative evaluation data and indicating areas for improvement.

8.6.1. Comparison of results of heuristic evaluation

The purpose of the *LSF* prototype was to apply VR technology to visualise generic hazards in the underground workplace, while the *ISGC* prototype specifically focused on the geological hazards that may contribute towards fall of ground if not managed correctly. In both prototypes, trainees had to identify the hazards and indicate how to address them safely. Computer-generated imagery was used to display the potential consequences of ignoring or not correctly addressing the hazards. In *ISGC,* trainees also had to identify the risks associated with the particular geological condition, prior to selecting control measures to remedy the situation. Due to the high fidelity required for distinguishing between the various geological conditions, more advanced graphics were used in the development of *ISGC* than those that were required for *LSF.*

To facilitate comparison between the evaluation response patterns of experts on the two training applications (*LSF* and *ISGC*), the frequencies of recorded evaluation rating scores of the experts on these two applications (condensed into 1-3 and 4-5 rating score categories) were cross-tabulated over the four heuristic categories of the study, namely *Instructional design* (Table 8.45), *General usability* (Table 8.46), *VR system design* (Table 8.47) and *Context-specific heuristics* (Table 8.49). The frequency counts reported in these tables reflect the totaled frequency-ratings of the evaluation statements for each of the eight heuristic criteria (the eight criteria are listed in each table). For example, in Table 8.45, the *Instructional design* category: for the criterion regarding *Clear goals, objectives or outcomes,* the frequency of 1 – 3 responses and 4 – 5 responses to the three sub-questions of this criterion for both the *LSF* and *ISGC* evaluations were tallied to yield the frequency patterns of 3 and 15; and 2 and 16 respectively. The rating levels of 1 to 3 and likewise 4 to 5 mentioned in this paragraph were condensed into two categories because it was argued that frequency counts would otherwise be too low for the reliable analysis of response patterns and the derivation of sound conclusions. It was argued that in the condensed rating scale, 1-3 would signify a perception of disagreement to indifference and a rating of 4 to 5 an extent of agreement.

Pearson's chi-square tests, and Cochran-Armitage trend tests (where applicable), were conducted on the frequencies of each of the four category tables, to compare the response patterns obtained in the *LSF* and *ISGC* evaluations. Because low cell-frequencies were expected in some frequency table cells (due to the sample size of six participating experts), the chi-square statistics calculated were compared to Fisher's exact probabilities for statistical significance. Fisher's calculation of exact probabilities, based on the hyper-geometric distribution, is more suited for these circumstances (Bower, 2003; SAS, 2014).

The subsections that follow compare the heuristic evaluation results of the *LSF* prototype to the heuristic evaluation results of the *ISGC* prototype. Each subsection contains:

- a table indicating rating-frequencies and chi-square values,
- a graph depicting frequency distributions of the rating frequencies,
- a graph showing average ratings per criterion, and
- discussion of the findings.

Category 1: Instructional Design

Table 8.45 shows the rating frequencies as explained in the previous section as well as the chi-square test statistic (with associated Fisher exact probability) calculated on the response distributions of the evaluation of *LSF* and *ISGC* for Category 1, Instructional Design (Tables 8.3 and 8.24 refer). The cells of the table also report the partial chi-square contribution of each cell frequency to the chi-square statistic.

The rating frequencies were determined by combining the ratings of all the evaluation statements per criterion, as indicated in Tables 8.3 and 8.24. These ratings were combined into two groups, one for all the 1, 2 and 3 ratings received (the more negative ratings) and one for all ratings of 4 and 5 (the positive ratings). For example, as indicated in Table 8.3, the number of 1, 2 and 3 responses received for all the evaluation statements for Criterion 1 in Category 1 is 3, while there were 15 responses of a 4 or 5 rating. Therefore, in Table 8.45 the frequency rating for Criterion 1 (Clear goals, objectives or outcomes) for *LSF* 1, 2 or 3 ratings is indicated as 3. Likewise, the frequency rating for Criterion 1 for *LSF* 4 or 5 ratings is indicated as 15. The combination of ratings into two groups was necessary because chi-square tests cannot be calculated for frequency tables where cell-frequencies for entire rows or columns are empty (=0), as was the case in many criteria for which a 1 rating was never assigned.

Once the rating frequencies were determined it was possible to conduct chi-square tests to establish whether the way evaluators responded to an evaluation statement differed from the way they responded to the other evaluation statements. The chi-square value (referred to as a chi-square statistic) is indicated in Table 8.45 underneath the rating frequency for each cell. The last row in Table 8.45 indicates Fisher's exact probability (0.039) associated with the particular chi-square statistic of 33.30 in the table.

Table 8.45: Criteria rating frequencies and chi-square statistic for Category 1.

Criteria	Rating frequency and partial chi-square contributions				Total
	LSF (1-3)	**LSF (4-5)**	**ISGC (1-3)**	**ISGC (4-5)**	
1. Clear goals, objectives or outcomes	3 0.0037	15 0.0008	2 0.9733	16 0.2749	36
2. Instructional assessment	2 0.0025	10 0.0005	2 0.1564	10 0.0442	24
3. Feedback to user responses	5 0.1773	19 0.037	5 0.0154	19 0.0044	48
4. Motivation and creativity	1 3.3717	29 0.7034	4 1.0288	26 0.2906	60
5. Differences between individual users	6 2.6933	12 0.5619	9 6.3967	9 1.8067	36
6. Reduction of extraneous processing	4 0.2566	14 0.0535	4 0.0003	14 0.0001	36
7. Fostering of germane cognitive load	8 1.5372	22 0.3207	11 2.9207	19 0.8249	60
8. Appropriate intrinsic cognitive load	0 3.1071	18 0.6483	0 3.9643	18 1.1197	36
Total	29	139	37	131	336
Fisher's exact probability associated with chi-square statistic of 33.30 is 0.039* Significance legend: significance on 5%, 1% and 0.1% is indicated as *; **; *** respectively					

The probability of 0.039 (which is less than the 5% significance level of 0.05) associated with the chi-square statistic of 33.30 in Table 8.45 indicates that evaluators did not respond in the same way to all eight criteria and that for the two prototypes they sometimes had statistically significantly different opinions regarding a specific criterion.

This is also shown in Figure 8.3, which graphically presents the rating frequencies per prototype group, as shown in Table 8.45. The grouping called *LSF negative* represents the number of ratings received for *LSF* ratings 1, 2 and 3 combined (indicated in blue), *LSF positive* represent the number of 4 and 5 ratings received for *LSF* (indicated in red). Similarly, the *ISGC negative* and *ISGC positive* groupings represent the same rating frequency groups for the *ISGC* prototype, indicated in green and purple respectively. The heuristic evaluation questionnaire used a five-point scale, with 1 representing *strongly* *disagree* and 5 *strongly agree*. Combining ratings 1, 2 and 3 to represent *negative* ratings therefore means that it includes *strongly disagree*, *disagree* and *neutral* ratings, whereas the positive grouping consists of *agree* and *strongly agree* ratings.

The preferred scenario in the graph in Figure 8.3 would be high frequencies on the red and purple bars, indicating positive responses, and low frequencies on the blue and green bars for negative responses, as is indeed mostly the case in the figure. In this regard, Figure 8.3 shows that the response patterns to Criteria 5 and 8 are different from the others. This deduction is verified by examining the larger cell-chi-square contribution values in these instances – the second entry in each cell of Table 8.45.

For Criterion 5 (Differences between individual users), the differentiation between the evaluation of the two tests is noticeable: a more positive evaluation was reported for the *LSF* test on this criterion and a neutral evaluation on the *ISGC* evaluation (the ratios of the blue and red bars to the ratios of the green and purple bars). Table 8.45 reports six 1-3 and twelve 4-5 responses when evaluating *LSF,* while there are nine 2-3 and nine 4-5 responses respectively for *ISGC*.

The response pattern for Criterion 8 (Appropriate intrinsic cognitive load) also indicates a significantly different response pattern from the other criteria: for both prototypes only positive responses were reported. It can be concluded that evaluation of all criteria (except Criterion 5 for the *ISGC* prototype) were positive.

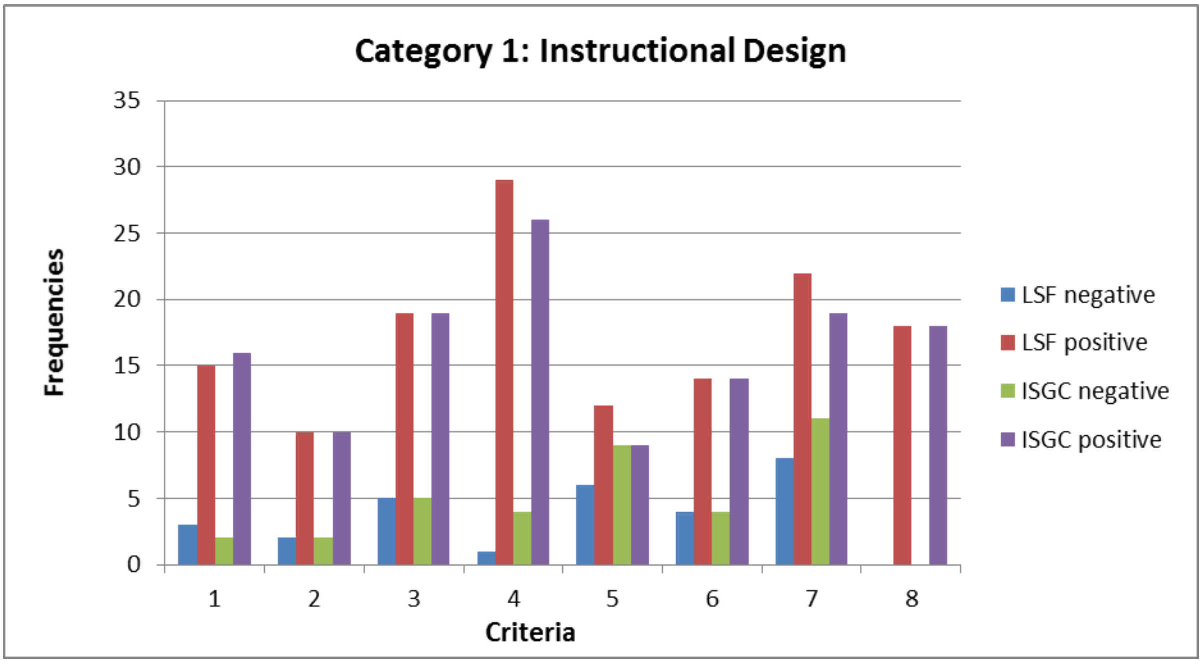


Figure 8.3: Frequency distribution of *LSF* and *ISGC* evaluation ratings for criteria in Category 1.

The deduction of a positive evaluation of Category 1 on all criteria for both tests is illustrated in Figure 8.4 which compares the criterion averages of Category 1 for the two prototypes. *LSF* mean evaluations are indicated in red and *ISGC* evaluations in blue. The means of all bars exceed 3.5, which signifies an extent of agreement or a positive evaluation. The average for each criterion for each test was calculated by totalling the products of each rating level (1 – 5) and frequency of occurrence of the rating level, and then dividing the total by the number of responses to evaluation statements for a specific criterion. For example, from Table 8.3, responses to the three evaluation statements were one 3 rating, two 2 ratings, four 5 ratings and eleven 4 ratings, and there were eighteen responses in total received for the three evaluation statements. The *LSF* Criterion 1 mean is thus calculated as 1x3 + 2x2 + 4x5 + 11x4) / 18 = 3.94.

Figure 8.4 indicates that definite positive evaluation of both prototypes were expressed (mean scores greater than 4) on criteria 2, 3, 4, 6 and 8; as well as Criterion 1 for *ISGC*. Criteria 5 and 7 were regarded as slightly less positive (with mean scores raging between 3.5 and 4). *LSF* clearly performed better than *ISGC* in this category and received higher average ratings than *ISGC* for all criteria except the first two.

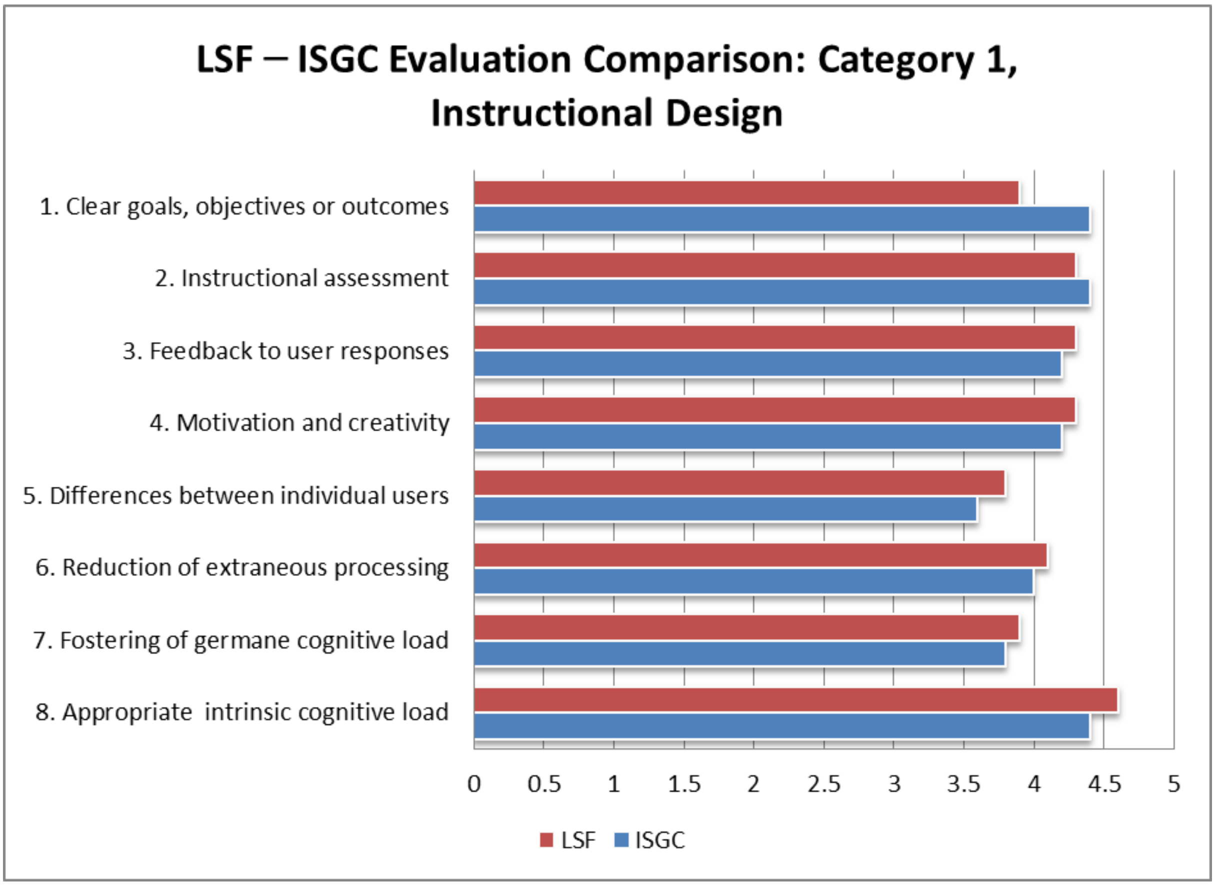


Figure 8.4: Comparison of average ratings in the *LSF* and *ISGC* evaluations for criteria in Category 1.

In the following subsections, similarly to Figure 8.4, Figures 8.6, 8.9 and 8.11 visually display comparisons between *LSF* and *ISGC* criteria averages for Categories 2, 3 and 4 respectively, each time using red to indicate the *LSF* criteria and blue for the *ISGC* criteria.

Category 2: General Usability

Similarly to the format used in Table 8.45, Table 8.46 shows the rating frequencies and chi-square test statistic (with associated exact probability) for responses received in the evaluation of *LSF* and *ISGC* for Category 2.

Table 8.46: Criteria rating frequencies and chi-square statistic for Category 2.

Criteria	**Rating frequency and partial chi-square contributions**				**Total**
	LSF (1-3)	**LSF (4-5)**	**ISGC (1-3)**	**ISGC (4-5)**	
1. Functionality	5 0.119	13 0.0372	5 0.9921	13 0.2157	36
2. User guidance	14 12.014	10 3.7545	11 10.519	13 2.2867	48
3. Consistency	1 3.8893	23 1.2154	0 4.2857	24 0.9317	48
4. Error correction	6 0.6857	12 0.2143	6 2.4143	12 0.5248	36
5. System status	3 0.3857	15 0.1205	3 0.0143	15 0.0031	36
6. Aesthetics	5 0.0893	19 0.0279	3 0.3857	21 0.0839	48
7. Error prevention	5 0.0893	19 0.0279	1 2.519	23 0.5476	48
8. Interactivity	1 2.519	17 0.7872	1 1.5254	17 0.3316	36
Total	40	128	30	138	336
Fisher's exact probability associated with chi-square statistic of 53.56 is less than 0.0001*** Significance legend: significance on 5%, 1% and 0.1% is indicated as *; **; *** respectively					

The exact probability (< 0.001) associated with the chi-square test statistic of 53.56 indicates statistical significance on the 0.1% level of significance. From Table 8.46 it can therefore be deduced that the response patterns to some criteria were significantly different from others on the 0.1% level of significance (Probability (chi-square statistic value being 53.56 under the null-hypothesis of no difference in response patterns over the criteria) is < 0.001). By looking at the larger chi-square values, which are the second entries in each cell of the table, we can detect that the response patterns to

Criteria 2 (User guidance), 3 (Consistency) and 8 (Interactivity) seems different from the others. This can also be seen in Figure 8.5, where the frequency distribution of the ratings is indicated graphically.

For Criterion 2, participants seem to exhibit an inverse satisfaction pattern to that which they exhibit for the other seven criteria for both prototypes: fourteen 1-3 responses and ten 4-5 responses for *LSF* – this is a more negative perception as opposed to eleven 1-3 responses and thirteen 4-5 responses for *ISGC*, which is somewhat positive but still neutral/undecided. For Criteria 3 and 8, participants indicated a slightly higher degree of satisfaction than for the other five remaining criteria.

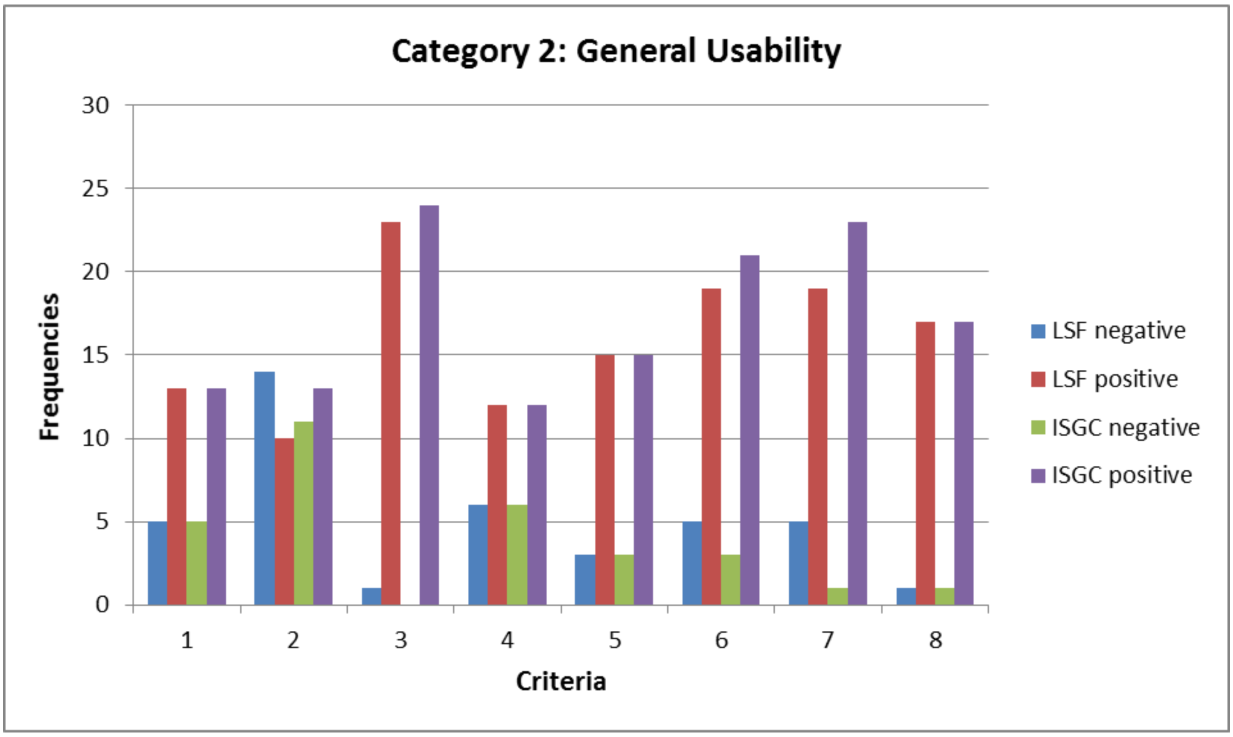


Figure 8.5: Frequency distribution of *LSF* and *ISGC* evaluation ratings for criteria in Category 2.

From the evaluation averages graph in Figure 8.6, it is noticeable that, for both prototypes, Criterion 2 (User guidance) received the lowest ratings, with *ISGC* scoring slightly higher than *LSF* (3.6 vs 3.2). Two other criteria also show notable differences between the two prototypes' assigned average ratings:

- Criterion 6 (Aesthetics): the evaluators much preferred the aesthetics of the *ISGC* prototype and assigned an average rating of 4.6, whereas *LSF* only scored 4.1 for this criterion.
- Criterion 7 (Error prevention): *ISGC* performed better in this criterion and scored an average 4.8 versus 4.3 for *LSF*.

High scores were achieved in Criterion 3 (Consistency) for both prototypes, while they scored exactly the same average for Criterion 8 (Interactivity). *ISGC* achieved higher average ratings than *LSF* in Criteria 2, 4, 6 and 7, while *LSF* was evaluated more positively than *ISGC* in Criteria 1, 3 and 5.

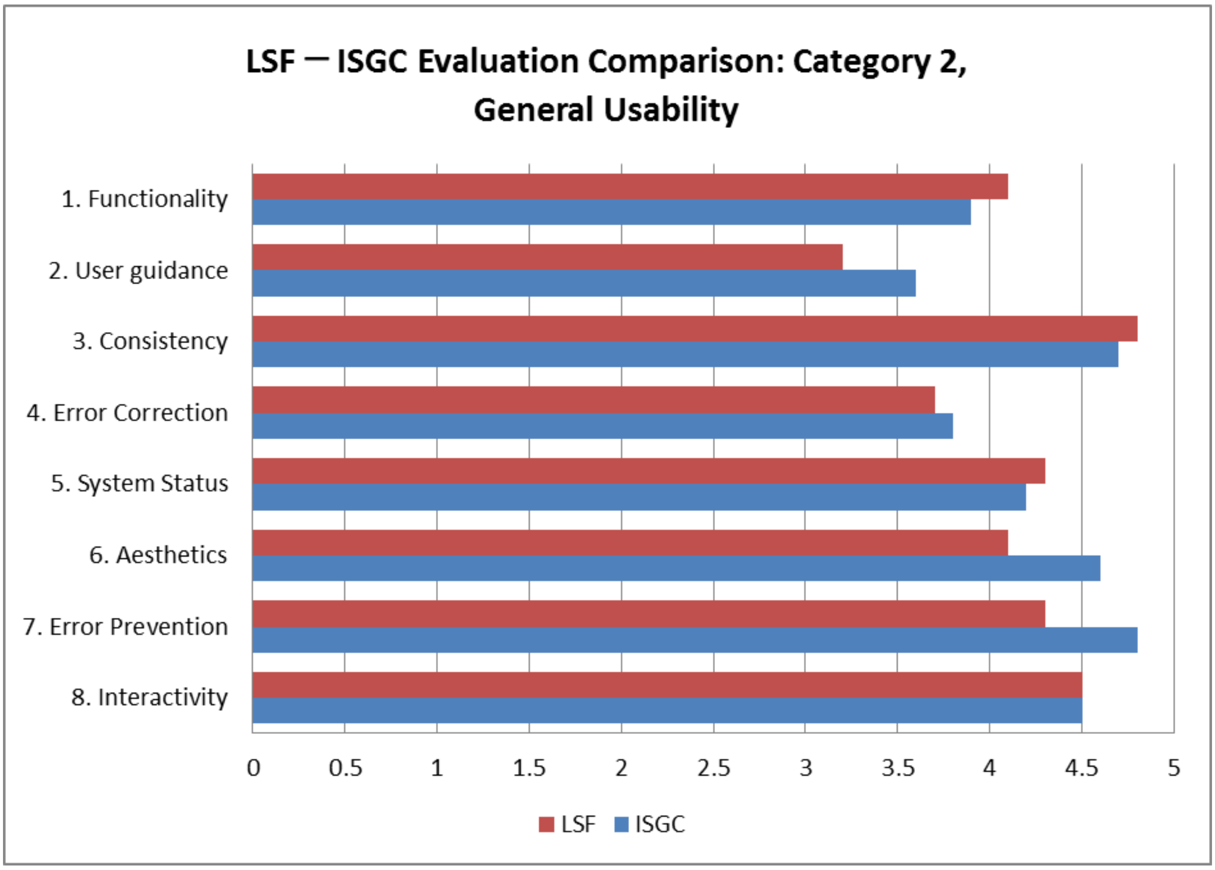


Figure 8.6: Comparison of average ratings in the *LSF* and *ISGC* evaluations for criteria in Category 2.

Category 3: Virtual Reality System Design

Table 8.47 indicates the rating frequencies and chi-square test statistic (with associated exact probability) for responses received in the evaluation of *LSF* and *ISGC* for Category 3. The chi-square test indicated that the response patterns to some criteria were statistically significantly different from other response patterns on the 0.1% level of significance (Probability of chi-square test statistic assuming a value of 61.37 under the null-hypothesis of no difference in response patterns for the eight criteria is < 0.001). This can be abbreviated to Probability (chi-sq = 61.37) < 0.0001***.

Table 8.47: Criteria rating frequencies and chi-square statistic for Category 3.

Criteria	Rating frequency and partial chi-square contributions				Total
	LSF (1-3)	LSF (4-5)	ISGC (1-3)	ISGC (4-5)	
1. User control	14 2.0167	10 1.3444	15 0.583	9 0.6149	48
2. Multimodal system output / feedback	6 3	24 2	9 2.6597	21 2.8055	60
3. Presence	2 1.6333	10 1.0889	3 1.621	9 1.7099	24
4. Orientation	3 2.45	15 1.6333	6 1.1361	12 1.1984	36
5. Navigation	4 0.1333	8 0.0889	5 0.2184	7 0.2304	24
6. Object interaction: selection and manipulation	10 1.0889	8 0.7259	10 0.0625	8 0.0659	36
7. Fidelity	9 0.0375	15 0.025	18 2.6187	6 2.7622	48
8. Various user modes	12 10.8	0 7.2	11 3.8029	1 4.0112	24
Total	60	90	77	73	300
Fisher's exact probability associated with chi-square statistic of 61.37 is > 0.001*** Significance legend: significance on 5%, 1% and 0.1% is indicated as *; **; *** respectively					

In general, participants evaluated Criteria 2-5 positively for both prototypes and negatively on Criteria 1, 6 and 8. Opposing views were expressed on Criterion 7 for the two prototypes. These views are illustrated in Figure 8.7.

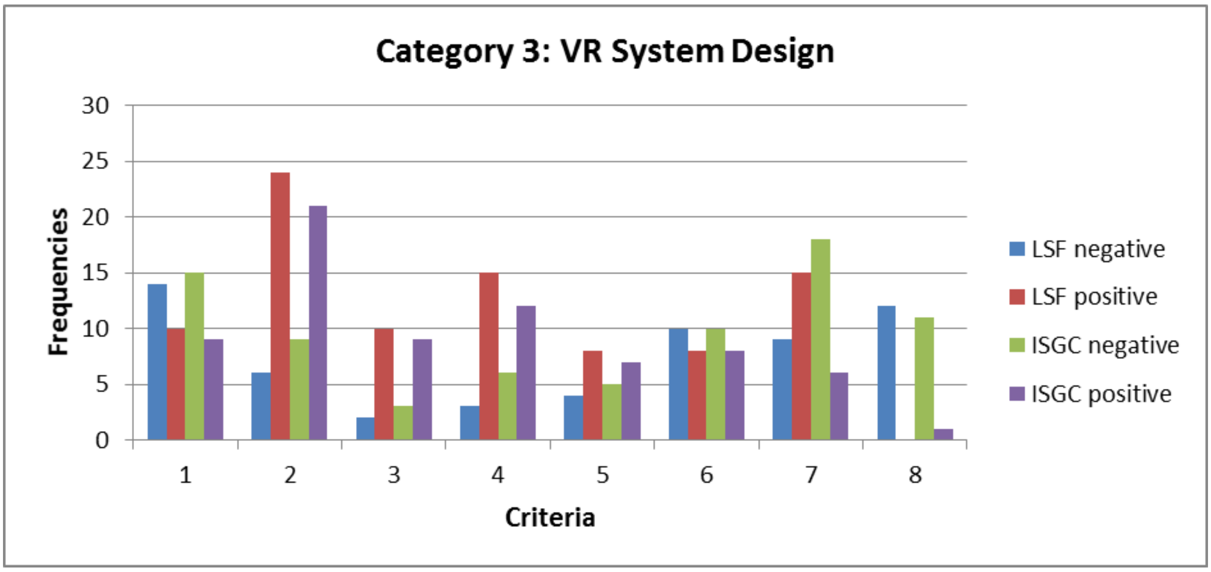


Figure 8.7: Frequency distribution of *LSF* and *ISGC* evaluation ratings for criteria in Category 3.

The nature of the differences in opinion ratings for Criterion 7 (nine 1-3 and fifteen 4-5 responses for *LSF,* and eighteen 1-3 and five 4-5 responses for *ISGC*) was further investigated by means of a Cochran-Armitage trend test. This test examined the refined agreement rating frequency distribution (levels 1-5 and not the condensed 1-3 and 4-5 levels) for the two prototypes. Results reported in Table 8.48 indicate a statistically significant difference (on the 0.1% significance level) in agreement response pattern for the two prototypes: the trend was stronger disagreement – a negative response – to the *ISGC* prototype and a tendency to positively rate the *LSF* prototype. (Probability (Z=3.78) < 0.001)). The trend is illustrated in the bar graph in Figure 8.8.

Table 8.48: Detail frequency evaluation ratings of the two prototypes and Cochran-Armitage trend test statistic for Criterion 7 in Category 3.

Likert rating scale value	**Rating frequency and column percentages**		**Total**
	LSF	**ISGC**	
1. Strongly disagree	0 (0.00)	5 (20.83)	5
2. Disagree	3 (12.50)	5 (20.83)	8
3. Neutral	6 (25.00)	8 (33.33)	14
4. Agree	2 (8.33)	5 (20.83)	7
5. Strongly agree	13 (54.17)	1 (8.34)	14
Total	24 (100.00)	24 (100.00)	48
The probability of the Cochran-Armitage statistic, the Z statistic assuming the value of 3.75 is less than 0.001 (62.5% *LSF*-agreement to 73.99% *ISGC* neutral to disagree perceptions)			

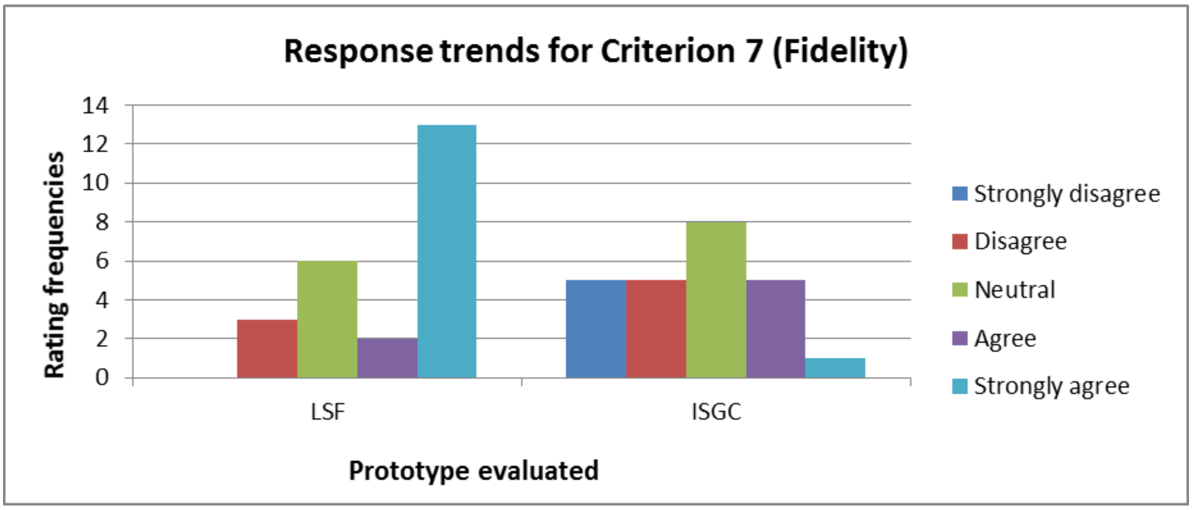


Figure 8.8: Different response trends for *LSF* and *ISGC* evaluation ratings for Criterion 7 in Category 3.

Figure 8.9 graphically displays the comparison between the Category 3 results of the two prototypes. From this graph it is clear that both *LSF* and *ISGC* were rated much lower for Criterion 8 (Various user modes) than for the other criteria, with Criterion 8 scoring 1.7 for *LSF* and 2 for *ISGC*. The graph also clearly indicates that *ISGC* outperformed *LSF* on all criteria in this category, except for Criterion 1 where both received the same average rating of 2.9 – which expressed a rather neutral evaluation.

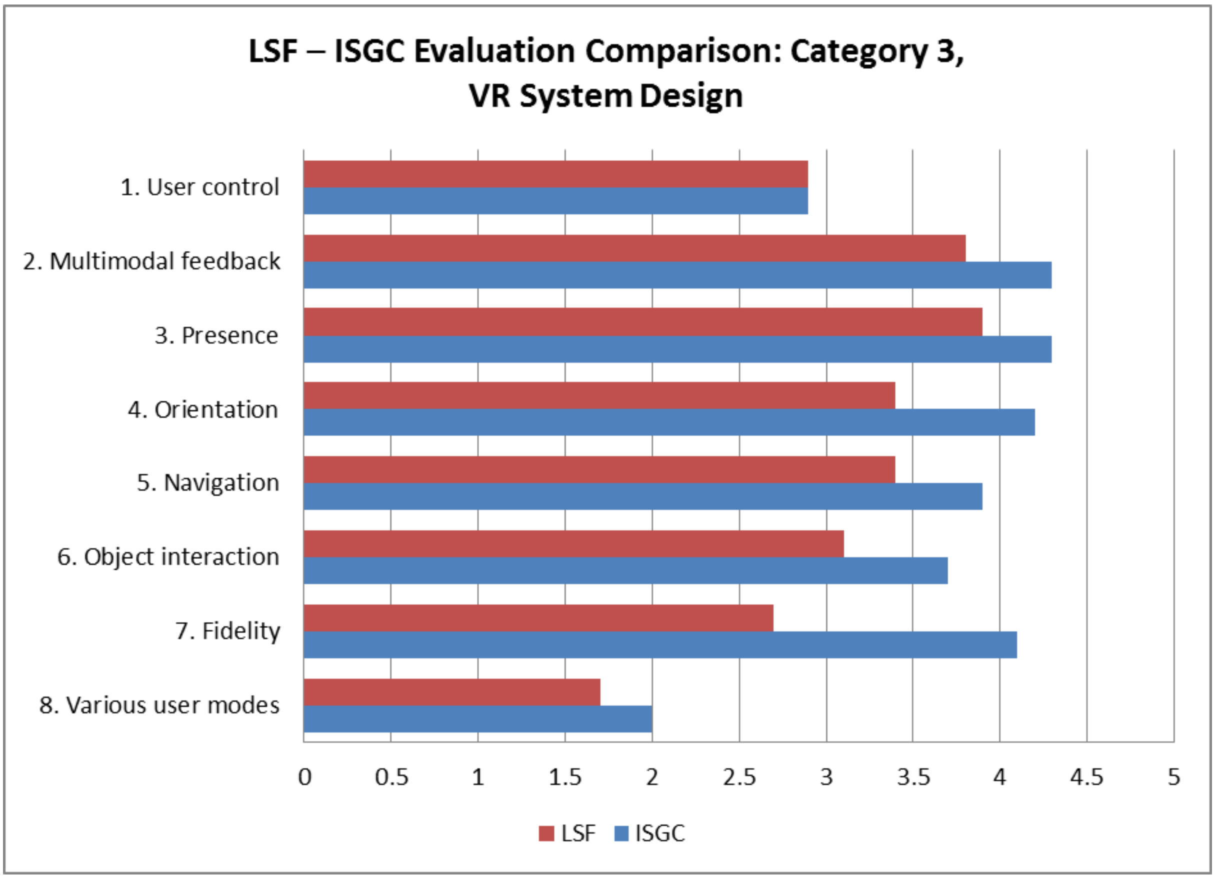


Figure 8.9: Comparison of average ratings in the *LSF* and *ISGC* evaluations for criteria in Category 3.

Only criteria 2, 3, 4 and 7 for *ISGC* had mean scores greater than 4 and on all criteria exceeded *LSF*. Criterion 8 was rated poorly for both prototypes.

Category 4: Context-specific heuristics

Similar to the discussion of the evaluation of the previous three categories, Table 8.49 indicates the rating frequencies and chi-square values for responses received in the evaluation of *LSF* and *ISGC* for Category 4, and Figure 8.10 displays the frequency distribution of the ratings graphically.

Table 8.49: Criteria rating frequencies and chi-square values for Category 4.

Criteria	Rating frequency and chi-square values				Total
	LSF (1-3)	**LSF (4-5)**	**ISGC (1-3)**	**ISGC (4-5)**	
1. Authentic tasks	1 2.4853	17 0.7682	4 0.1731	14 0.0381	36
2. Appropriate reference materials	17 38.25	1 11.823	9 10.173	9 2.2415	36
3. Comprehensive scope	2 0.2451	10 0.0758	2 0.0128	10 0.0028	24
4. Adaptive design	6 0.7206	12 0.2227	1 1.5577	17 0.3432	36
5. Appropriate record keeping	0 4.25	18 1.3136	4 0.1731	14 0.0381	36
6. Trainee preparedness	2 0.2451	10 0.0758	1 0.6282	11 0.1384	24
7. Relevant subject matter	1 5.2245	29 1.6148	3 1.0782	27 0.2376	60
8. Understandable and meaningful symbolic representation	5 0.1324	13 0.0409	2 0.4808	16 0.1059	36
Total	34	110	26	118	288
Fisher's exact probability associated with chi-square value of 84.91 is less than 0.001*** Significance legend: significance on 5%, 1% and 0.1% is indicated as *; **; *** respectively					

The chi-square test indicated that the response patterns to some criteria were statistically significantly different from other response patterns on the 0.1% level of significance (Probability (chi-sq = 84.91) < 0.001). The participants, in general, exhibited a positive perception for all criteria for both prototypes, except for Criterion 2 (Appropriate reference materials), which was negatively evaluated for *LSF* and neutral for *ISGC*.

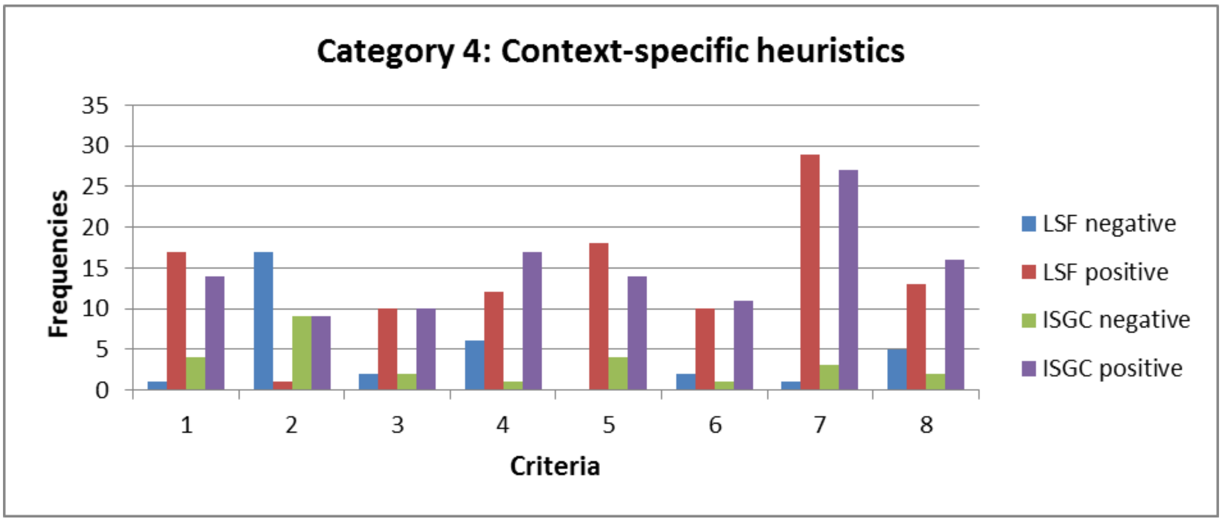


Figure 8.10: Frequency distribution of *LSF* and *ISGC* evaluation ratings for criteria in Category 4.

The comparative average scores for Category 4 are shown in Figure 8.11. The only criterion where the two prototypes scored equally is Criterion 1 (Authentic tasks) where both scored 4.3. On all other criteria (except Criterion 5) *ISGC* was evaluated more favourably than *LSF*. For Criterion 5, the mean *LSF* rating was 4.7 and that of *ISGC* 4.5. For Criterion 2 (Reference materials), *LSF* was rated very poorly with a mean score of 1.7 (and 2.8 for the *ISGC* prototype). *ISGC* scored an average of 4.7 on Criterion 6 (User preparedness). The scores for Criterion 2 for both prototypes are noticeably lower than the other criteria. The reasons for this were highlighted in the discussion of the detailed results in Table 8.27.

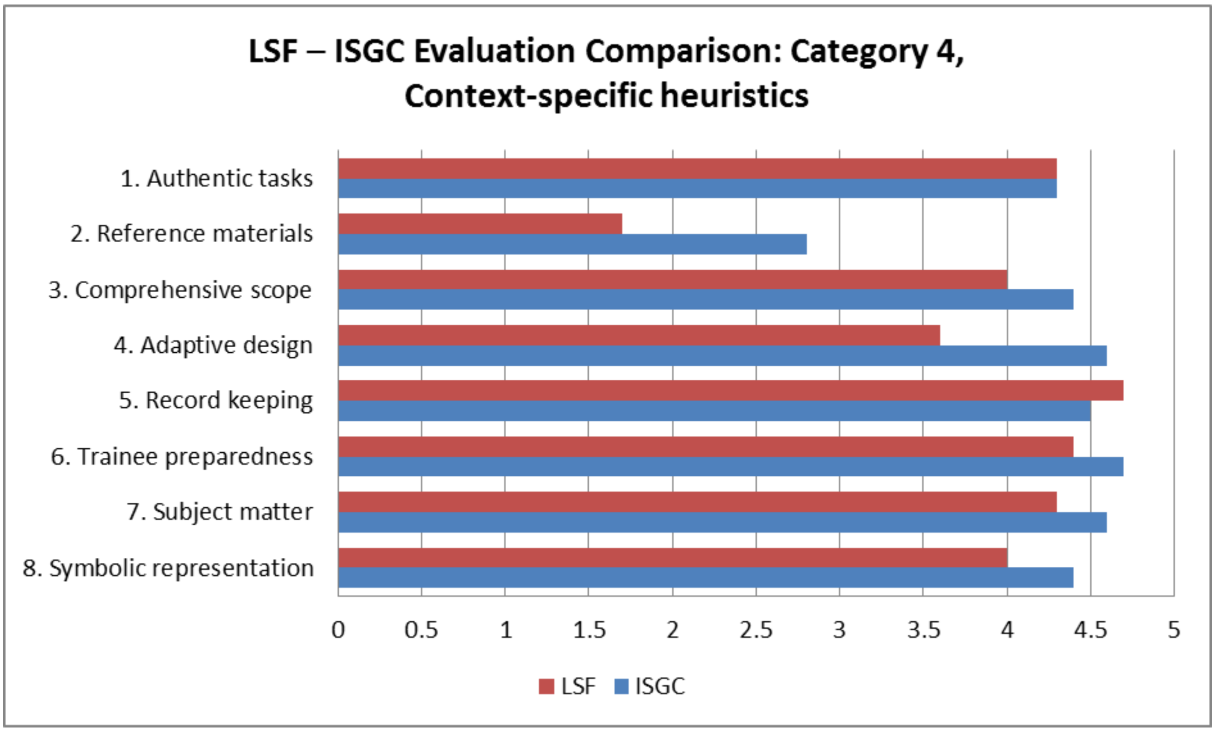


Figure 8.11: Comparison of average ratings in the *LSF* and *ISGC* evaluations for criteria in Category 4.

Discussion of comparative findings

A graph depicting the comparative average ratings for all four categories is shown in Figure 8.12. The average for each category is indicated next to the relevant category bar. This was calculated as the total of the six evaluator averages divided by six.

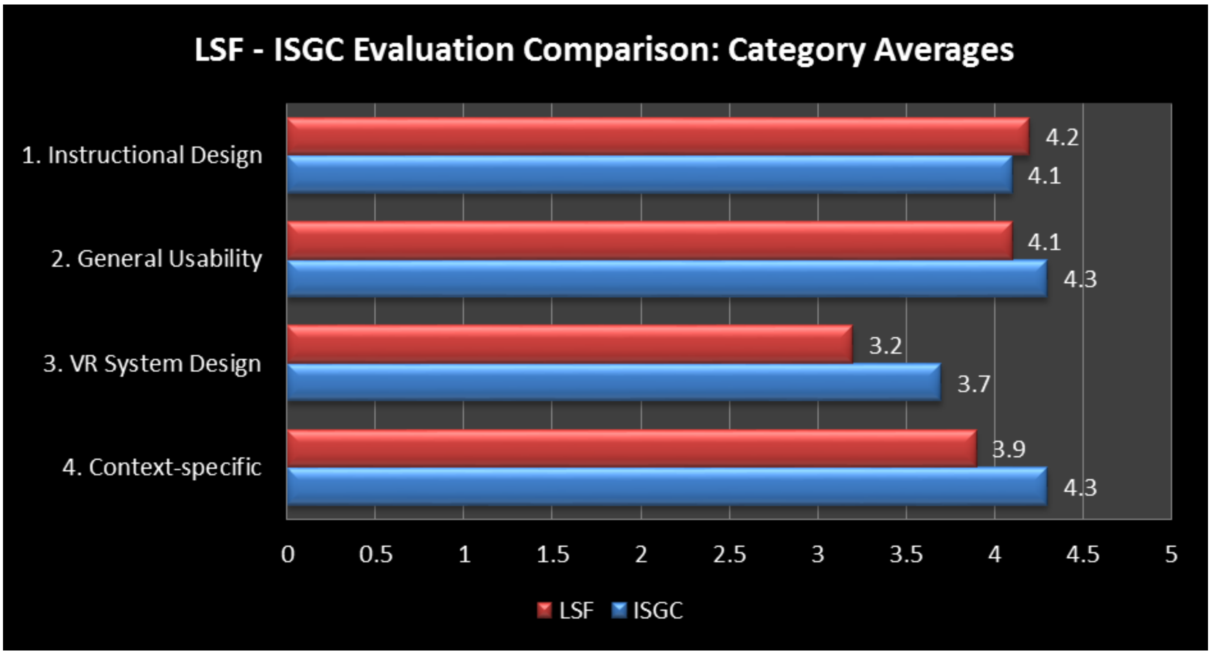


Figure 8.12: Comparison of *LSF* and *ISGC* evaluation averages per category.

It is evident that *LSF* achieved a slightly higher average than *ISGC* for Category 1, but *ISGC* scored better in all three of the other categories. The averages do not differ significantly between the two prototypes, with the highest difference being 0.5 for Category 3 (3.2 for *LSF* vs 3.7 for *ISGC*). The best category for *LSF* was Category 1, with an average of 4.2. *ISGC* had two highest averages of 4.3, that is, for Category 2 and Category 4.

Both prototypes received their lowest average ratings for Category 3 – VR Systems Design. This is important to note, as the intention of the development of the prototypes was to demonstrate how VR can be utilised to improve training, and it is specifically the VR features that the evaluators indicated should improve the most. The *DEVREF* Framework performed well in both cases in supporting the emergence of these inadequacies.

ISGC was developed as an exploratory VR training prototype focusing on geological hazards and utilising alternative development tools to improve realism, and not as an improved version of the *LSF* prototype. The *ISGC* prototype did, however, incorporate the usability improvements identified during design reflection of the *LSF* prototype. The comparison of the evaluation results of the two prototypes indicates that *ISGC* improved on *LSF* in three of the four categories, but did not achieve a higher rating than *LSF* in instructional design. This can be ascribed to the fact that this was the first effort to develop a VR training prototype specific to geological hazards. Moreover, the instructional design of *ISGC* was to a large extent a new design and not an extension of the *LSF* design.

The other important distinguishing factor between *ISGC* and *LSF* is that *ISGC* required high-fidelity computer-generated 3D graphics to portray the geological conditions realistically. This necessitated the use of more advanced development tools, not only influencing fidelity, but also other aspects of the VR system design, such as presence, orientation, navigation and object interaction. The comparison of the evaluation results of the two prototypes demonstrates that *ISGC* improved on the *LSF* evaluation rating for the VR system design aspects, evaluated in Category 3, even though these development tools were used for the first time.

8.6.2. Comparison of results of the user satisfaction questionnaires

This section compares the results of the *LSF* user satisfaction evaluation with the results of the *ISGC* user satisfaction evaluation. The same questionnaire was used for both prototypes, but since the sizes of the two user groups were significantly different (195 versus 52), the comparison results are presented as percentages of each group.

Comparison of findings for biographic details

Table 8.50 indicates the educational levels of participants in the two studies. Distribution of the educational levels of the two groups is similar, with 57% of the *LSF* participants and 58% of the *ISGC* participants having completed secondary school (Grade 12). Similarly, 3% of *LSF* participants and 4% of *ISGC* participants had left primary school before completing Grade 6. This indicates that those who had been promoted to the higher ranks (*ISGC* participants), had been promoted due to performance and expertise and not due to higher educational levels.

Table 8.50: Percentage comparison of LSF and ISGC participants' schooling levels.

Schooling level	LSF	ISGC
< Grade 6	3%	4%
Grade 6-7	4%	8%
Grade 8-9	8%	9%
Grade 10-11	28%	21%
Grade 12	57%	58%

Figure 8.13 compares the participants' exposure to technological devices prior to using the system, as indicated in Tables 8.11 and 8.32 respectively. It is clear that the *ISGC* users, as a group, had considerably more exposure to the use of computers (71%) than the *LSF* group (34%), although they had similar exposure to cell phones and ATMs.

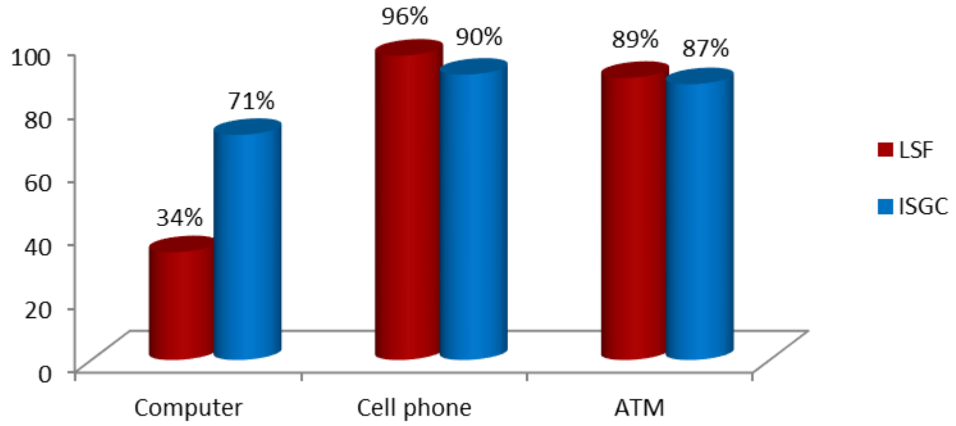


Figure 8.13: Percentage comparison of *LSF* and *ISGC* participants' previous technological device exposure.

Prior to using the prototypes, participants had the option of doing pre-training to prepare them for the required interaction. The training mainly entailed use of the computer mouse. Figure 8.14 indicates what percentage of participants opted to do the pre-training. Fewer *ISGC* participants did the pre-training (62%) than *LSF* participants (85%), which is in line with the information in Figure 8.13 that *ISGC* participants had more prior exposure to computers than the *LSF* group.

Figure 8.14: Percentage comparison of *LSF* and *ISGC* participants choosing to do pre-training.

The pie charts in Figure 8.15 compare the selected system language of the two groups of participants. As explained in Section 8.4.1.1, both prototypes catered for English, Xhosa, Tswana and Sepedi, and participants could choose their language of use. It is clear from Figure 8.15 that in the case of the *LSF* prototype, most participants (68%) chose to do the system in Tswana, whereas with the *ISGC* prototype the majority of participants (86%) chose English. This seems to suggest that in the first level of management (*ISGC* users), English is more prevalent than in the general workforce. Sepedi is not commonly used in the Rustenburg region, which explains why only two *LSF* participants (1%) chose Sepedi as the system language, while none of the *ISGC* participants chose Sepedi.

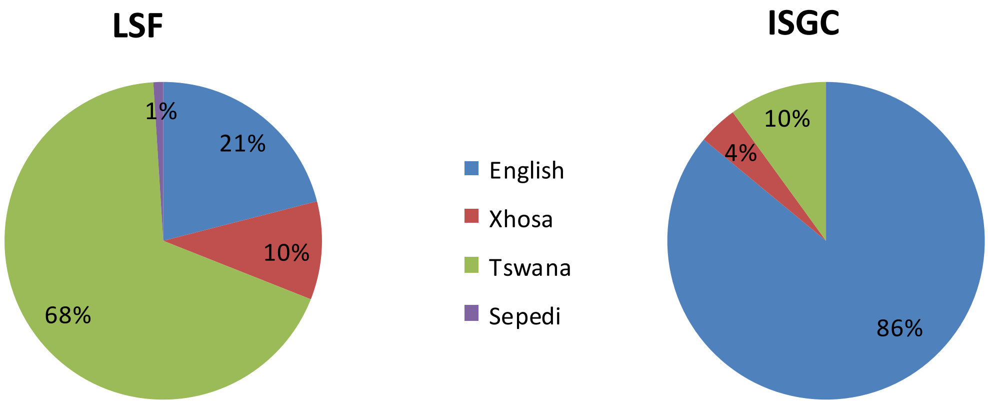


Figure 8.15: Percentage comparison of selected language for *LSF* and *ISGC*.

Table 8.51 gives a breakdown of the ages of all the participants for both prototypes. The *LSF* participants were aged between 20 and 48, while the *ISGC* participants' ages varied between 25 and 60. It is clear from the table that the majority of *LSF* participants were between 25 and 35 years of age. For *ISGC* participants, no age group was more prevalent or less prevalent than others, except that one participant was 60, while there were no participants between 53 and 60. The slightly higher ages of *ISGC* participants is due to the target group being from higher ranks of underground workers, who are more experienced and have been promoted to higher positions, as explained in Section 8.5.2.

Table 8.51: Age comparisons of the LSF and ISGC participants.

Age	LSF		ISGC	
	Frequency	Percentage	Frequency	Percentage
20 – 24	23	11.8	0	0
25 – 29	67	34.4	5	9.6
30 – 34	58	29.7	11	21.2
35 – 39	33	16.9	6	11.5
40 – 44	9	4.6	10	19.2
45 – 49	5	2.6	11	21.2
50 – 54	0	0	8	15.4
55 – 59	0	0	0	0
60+	0	0	1	1.9
Totals	195	100.0	52	100.0

__Comparison of findings for usability aspects__

The next six graphs represent the comparative findings for the five usability aspects: Ease of use, Learnability, Satisfaction, Authenticity (2 groups) and Method of choice, as discussed in Section 8.4.2.2 for *LSF* and Section 8.5.2.2 for *ISGC*. In each case, the graph represents the average percentage rating for all the questions within that specific usability category, for example, the *LSF* average shown in Figure 8.16 for a 1 rating for the Ease of use category, 57%, is the average percentage of the responses received for the six questions relevant to the Ease of use category, Questions 2.4, 2.5, 2.6, 2.7, 2.10 and 2.14. As indicated in Table 8.14, the percentages of responses of a 1 rating received for these questions were 44.6, 69.3, 42.5, 55.4, 64.1 and 64.6 respectively, of which the average is 340.5/6 = 57% of the 195 participants. All the other averages were calculated similarly for the questions relevant to that specific usability category, and the small percentages of 'no responses' were omitted from the graph data.

The same colour key is used in the following six graphs, presented as stacked column graphs. Dark blue represents the average percentage of all the 1 ratings received for the five questions in the category, red is used for the average percentage of all the 2 ratings, green for the 3 ratings, purple for the 4 ratings and light blue for the 5 ratings.

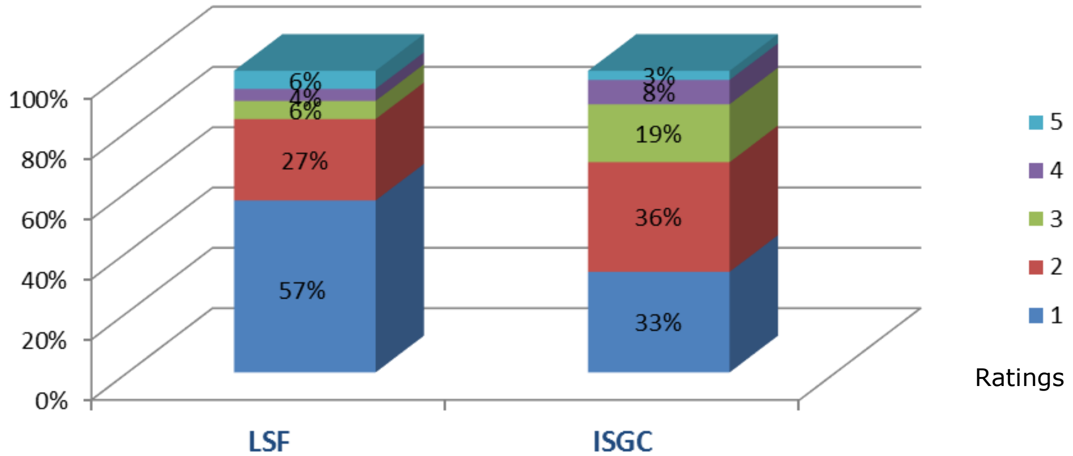


Figure 8.16: Percentage comparison of *LSF* and *ISGC* average ratings for the Ease of use category.

From the stacked column graph in Figure 8.16 it is clearly noticeable that the ease-of-use ratings received for *LSF* were better than the *ISGC* ratings, with 84% of the average *LSF* ratings being a 1 or 2, versus only 69% for *ISGC*. This is rather surprising, taking into account the higher computer exposure of the *ISGC* group as indicated in Figure 8.13, but it could point towards the success of the pre-training, which was done by 85% of the *LSF* participants and only 62% of the *ISGC* participants (as seen in Figure 8.14).

Comparing the two prototypes regarding the Learnability category, as shown in Figure 8.17, reveals that even though *LSF* received a higher average of 1 ratings than *ISGC* (74% versus 53%), there is not much difference between the total of the first two ratings (95% versus 92%), indicating a very high level of learnability for both systems, yet higher for *LSF*.

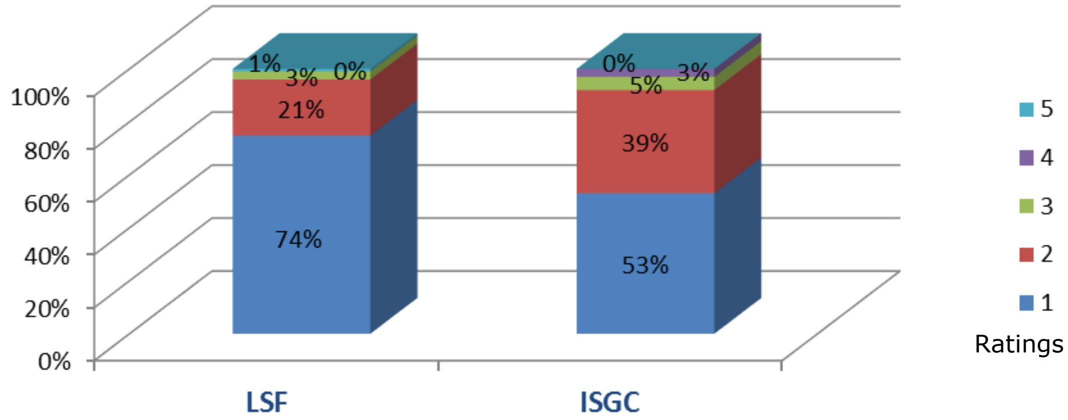


Figure 8.17: Percentage comparison of *LSF* and *ISGC* average ratings for the Learnability category.

Figure 8.18 presents the comparative average ratings for the Satisfaction category. Once again the average percentage of the 1 ratings is higher for *LSF*, but the 2 rating is higher for *ISGC*. Even though the total of the first two ratings (*Very satisfied* and *Much satisfied*) is 94% for *LSF* and only 87% for *ISGC*, these scores still represent a very high level of user satisfaction for both systems.

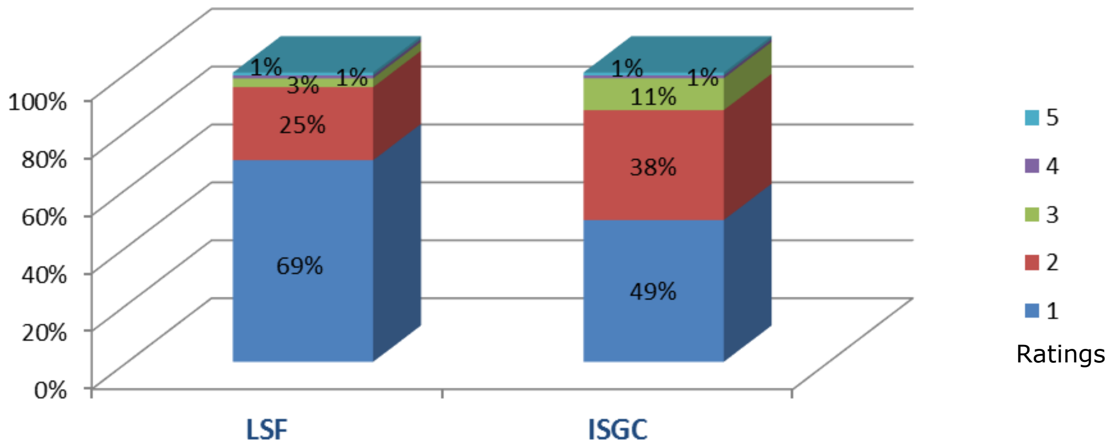


Figure 8.18: Percentage comparison of *LSF* and *ISGC* average ratings for the Satisfaction category.

As discussed in Section 8.4.2.2, the Authenticity category was divided into two subsections. The first subsection grouped the answers received to Questions 2.8, 2.13 and 2.17 together, since they were based on the same five-option Likert rating scheme as the questions covering the other usability traits. The answers received for Questions

2.11 and 2.12, however, were treated separately because they were based on a different four-option scale as an *Average* option was not feasible. The comparative results for these two subsections are presented in Figure 8.19 and Figure 8.20 respectively.

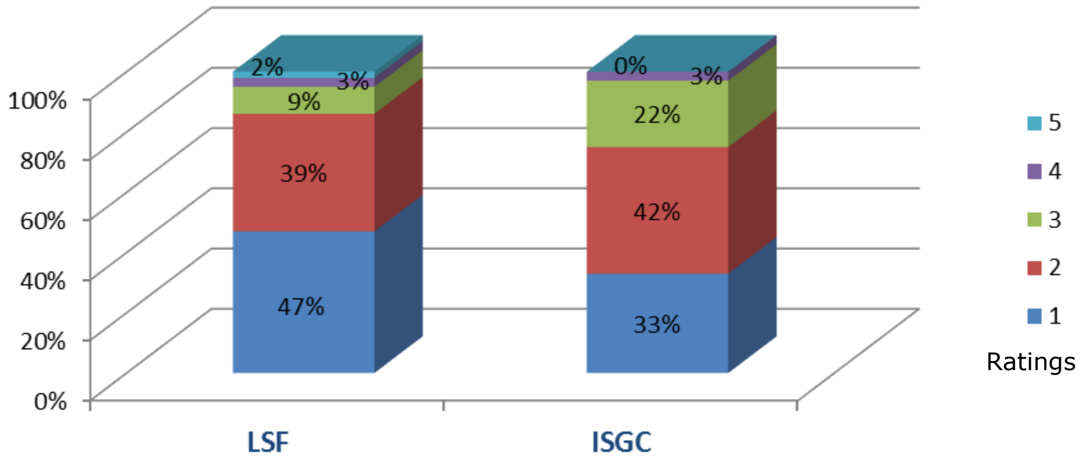


Figure 8.19: Percentage comparison of *LSF* and *ISGC* average ratings for the Authenticity category (Questions 2.8, 2.13 and 2.17).

Figure 8.19 indicates a very high authenticity average for both systems, with the *LSF* prototype again achieving higher ratings than the *ISGC* prototype. The *LSF* average percentage for the first two ratings is 86% and for the *ISGC* prototype it is 75%.

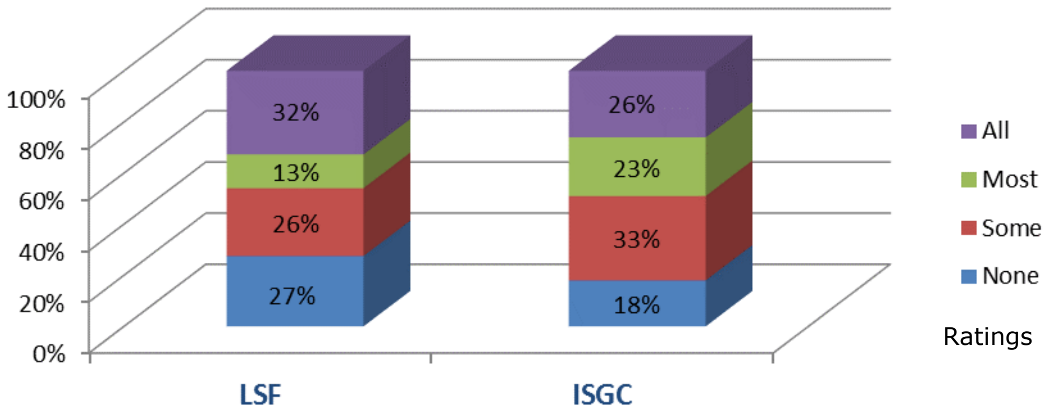


Figure 8.20: Percentage comparison of *LSF* and *ISGC* average ratings for the Authenticity category (Questions 2.11 and 2.12).

Question 2.11 dealt with the possibility of accidents occurring and Question 2.12 with the possibility of an accident happening to the participant personally. As can be seen in Figure 8.20, the group average percentages varied for all the options. In the *LSF* graph, 27%, on average, were of the opinion that the accidents could not happen, but in the case of *ISGC* this figure drops to 18%, indicating that this group has a better realisation of the danger of the hazards at their workplace. Nevertheless, the totals for *Most* and *All* responses in both groups, give fairly similar results with fewer than half of the respondents in both groups (45% of *LSF* participants and 49% of *ISGC* participants) believing that most or all of these accidents can actually occur at the Mine. This is a major concern, as the accidents portrayed in both systems were selected from actual incidents at the Mine.

Figure 8.21 indicates the average percentages for the Method of choice category, where a 1 rating represented a *Very high* preference for computer-based training similar to the prototypes, and a 2 rating represented a *High* preference. It is clear from the graph that the support for such computer-based training is indeed overwhelming, with 93% of the *LSF* group responses in the first two ratings, and also 88% of the *ISGC* group.

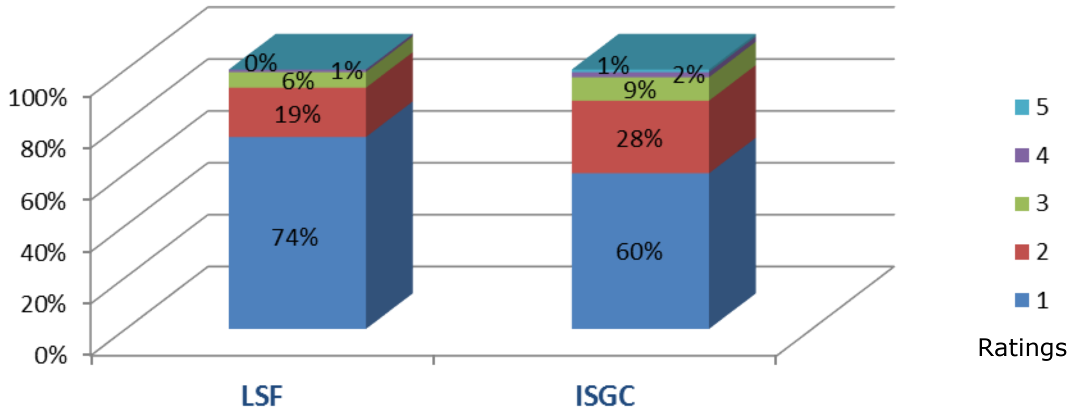


Figure 8.21: Percentage comparison of *LSF* and *ISGC* average ratings for the Method of choice category.

Discussion of comparative findings

As indicated in Table 8.50 and Figures 8.14 and 8.15, the educational levels of the participants in both groups are similar, but 71% of the *ISGC* participants had previous exposure to computers compared to the 34% of *LSF* participants. This led to only 62% of the *ISGC* participants doing the pre-training, while 85% of *LSF* participants opted to do it.

If we combine the average 1 ratings (e.g. *Very much*) with the average 2 ratings (e.g. *Much*), the comparison of the user satisfaction questionnaire results of the two prototypes indicated that *LSF* received slightly higher user ratings than *ISGC* for the five usability aspects depicted in the preceding six graphs, i.e. Ease of use, Learnability, Satisfaction, Authenticity and Method of choice. This could be due to a combination of factors:

- Fewer *ISGC* participants chose to do the pre-training and may therefore have struggled more than *LSF* participants with some of the usability aspects.

- The participants who evaluated *ISGC* held more senior ranks than the *LSF* participants, and might have expected more from the training system than the lower level workers.
- *ISGC* has a specialised focus on geological hazards with more complex content than the generic hazards depicted in *LSF*.

The Bonferroni multiple comparisons of means test conducted on the effect of schooling and pre-training on ease-of-use perceptions of *LSF* participants, as described in Section 8.4.2.5, indicated that higher schooling levels and pre-training contribute to the ease with which the training application is used. Similarly, a GLM analysis of the *ISGC* data, reported in Section 8.5.2.5, indicated that the participants with higher schooling levels found the system easier to use, but in this case, regardless of pre-training. These findings, however, only apply to the Ease of use category and do not detract from the overall positive evaluation of *LSF* and *ISGC*, as described in Sections 8.4.2 and 8.5.2, respectively.

In analysing the overall results of the evaluation of the two prototypes, it is evident that *ISGC* received higher ratings than *LSF* for the heuristic evaluation by experts, while *LSF* received slightly higher ratings than *ISGC* in the satisfaction questionnaire completed by end-users. This could be due to the following reasons:

- The *LSF* participants had notably lower prior computer exposure (33.8%) than the *ISGC* participants (71.2%), which may have resulted in them being more surprised and impressed with the training experience (the so-called 'wow factor'). This may have biased their ratings of some usability aspects.
- The user satisfaction questionnaire focused mostly on evaluating usability aspects, whereas the heuristic evaluation, applying the *DEVREF* evaluation instrument, focused on instructional design, VR system design and context-specific aspects as well. The scope of evaluation using *DEVREF* was therefore much more comprehensive and could result in differing findings overall.
- The HE evaluators are all experts in their respective fields and could evaluate technicalities and subtleties in much more detail than general users.

8.7. Conclusion

This chapter presented results of the empirical evaluations done on the *LSF* and *ISGC* prototypes, applying the *DEVREF* Evaluation Framework for the heuristic evaluations and the user satisfaction questionnaire for the user-based evaluations.

Sections 8.2 and 8.3 explained the design of the evaluation instruments used, while Section 8.4 discussed the evaluation of the *LSF* prototype by heuristic evaluation using the proposed *DEVREF* Evaluation Framework, and by a user satisfaction questionnaire. Section 8.5 reported on the evaluation of the *ISGC* prototype, also by applying the same *DEVREF* Framework and the user satisfaction questionnaire. The final section, Section 8.6, compared the evaluation findings of the two prototypes.

More important than the evaluations themselves, however, is the performance of the evaluation instrument based on the *DEVREF* Framework in order to underscore the value of such an evaluation framework. From the evaluation analysis in Sections 8.4.1 and 8.5.1 for the *LSF* and *ISGC* prototypes respectively, it is clear that the evaluation instrument based on the framework clearly revealed the usability and design aspects in which the prototypes performed well, and also aspects where usability improvements are required. Furthermore, the heuristic evaluations also identified a number of inadequacies in *DEVREF*, which prompted the researcher to validate the framework even further by performing a meta-evaluation on the framework. This meta-evaluation is described in detail in Chapter Nine.

Chapter Nine
Revised Evaluation Framework

9.1. Introduction

As indicated in Chapter Five, the design-based research process leads to dual outcomes: the development of theory, and a practical contribution in the form of a real-world innovative product or intervention. Chapters Six and Seven explained the design and evaluation of the practical contribution, namely the prototype systems *LSF* and *ISGC* respectively. These prototype systems have since become real-world training systems currently in use at various mines. This chapter discusses refinements to the theoretical contribution of this study, namely the *DEVREF* Evaluation Framework.

Section 5.5.4 in Chapter Five described the characteristics of DBR, and pointed out the views of various authors on the objective of developing new theory, including the following:

- DBR leads to the construction of theoretical frameworks that inform future designs (Bowler & Large, 2008).
- The research should lead to sharable theories that help communicate relevant implications to practitioners and other designers (DRC, 2003).
- DBR extends theories and refines design principles to eventually lead to substantial change in educational practice (Wang & Hannafin, 2005).

The *DEVREF* Framework, presented in Section 8.5, comprises four categories, covering the following aspects: instructional design, general usability of the design, virtual reality system design, and context-specific criteria. Although *DEVREF* is presented as an evaluation framework, the criteria in the framework also relate extensively to design aspects. Hence, *DEVREF* serves two purposes:

- It can be applied to evaluate existing interactive desktop VR training systems, and
- It can be applied to inform the design of such systems, as design principles are implicitly incorporated in the framework.

The outcome of the evaluations described in Chapter Eight not only provided valuable information regarding the *LSF* and *ISGC* prototypes, but also indicated that the evaluation framework itself had inadequacies. For this reason, a meta-evaluation of the *DEVREF* Framework was done to strengthen *DEVREF*. The meta-evaluation and its findings are discussed in this chapter, and address Research Subquestion 6 of this study regarding the appropriateness and effectiveness of the proposed framework. The evaluations of the *LSF* and *ISGC* prototypes, as well as the meta-evaluation, led to an improved version of the *DEVREF* Evaluation Framework, which is presented at the end of the chapter.

Figure 9.1 shows the topical layout of this chapter, indicating its various subsections. The coloured parts of the research process flow diagram in Figure 9.2 indicate that the entire DBR Cycle 4 is covered by this chapter. Section 9.2 deals with the concept of meta-evaluation and the design of the meta-evaluation questionnaire. Section 9.3 discusses the findings of the meta-evaluation relating to the framework's criteria and the framework's method. Based on the interpretation of all the findings presented, the researcher presents an improved version of the *DEVREF* Evaluation Framework in Section 9.4.

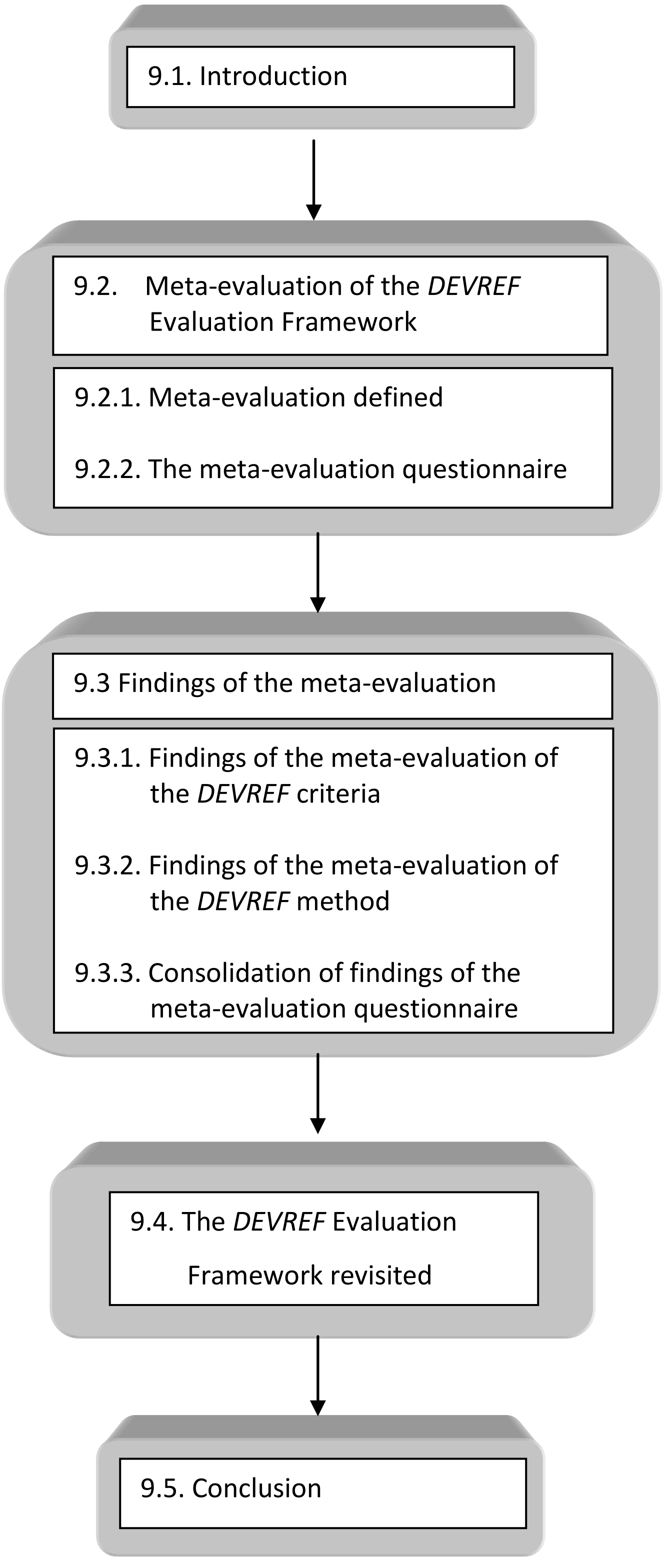


Figure 9.1: Layout of Chapter Nine.

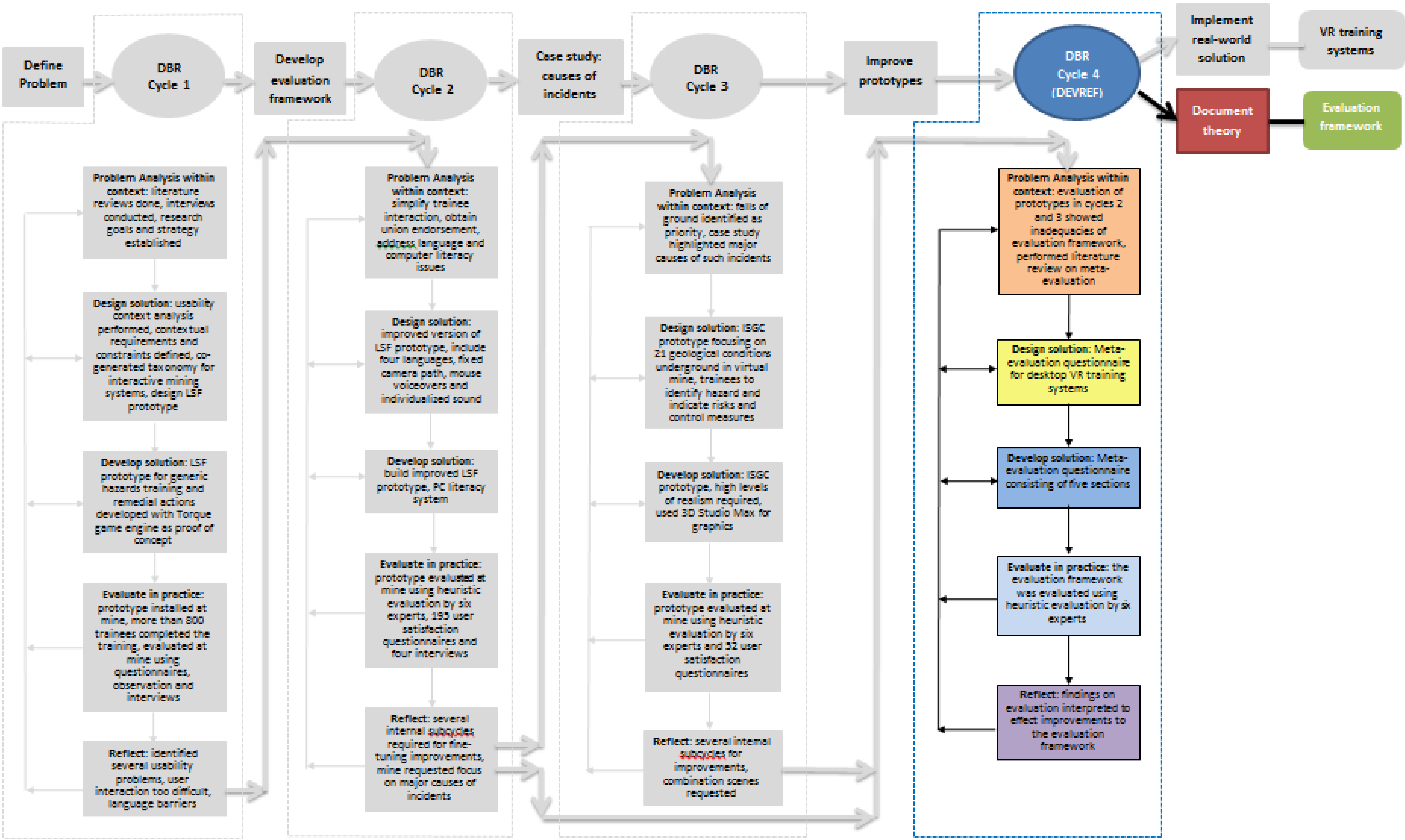


Figure 9.2: Research process flow discussed in Chapter Nine.

9.2. Meta-evaluation of the *DEVREF* Evaluation Framework

The next two sections address Research Subquestion 6 of this study:

RQ6	How appropriate and effective is the proposed framework?

Meta-evaluation is the evaluation of an evaluation. In compiling this meta-evaluation, the researcher undertook the following processes:

1. The literature available on meta-evaluation of evaluation frameworks was studied.
2. The literature used in generating the heuristic evaluation framework and criteria presented in Sections 3.3, 3.4 and 3.5, with the findings given in Section 5.8, was revisited.
3. Using the material studied for point 2, the most important aspects were consolidated into a set of evaluation statements to evaluate the *DEVREF* Evaluation Framework used in this research. This resulted in the meta-evaluation questionnaire, attached as Appendix B-3.
4. The meta-evaluation questionnaire was used to conduct the meta-evaluation with the same evaluators who had participated in the heuristic evaluation of the prototype VR training programs.
5. The findings of the meta-evaluation were interpreted and the framework revised accordingly.

9.2.1. Meta-evaluation defined

A limited amount of literature is available on meta-evaluation. A definition of meta-evaluation frequently referred to in the literature, is the one proposed by Stufflebeam (2001:183): "Meta-evaluation is the process of delineating, obtaining, and applying descriptive information and judgmental information about an evaluation's utility, feasibility, propriety, and accuracy and its systematic nature, competence, integrity/honesty, respectfulness, and social responsibility to guide the evaluation and publicly report its strengths and weaknesses". This comprehensive definition describes meta-evaluation as a process that investigates fine-grained detail of an evaluation in a comprehensive and meticulous manner. Stufflebeam (2011) expands this definition by adding that a meta-evaluation judges an evaluation against a set of ideas regarding what constitutes a good evaluation.

Meta-evaluation can be performed for different purposes, for example, as a tool to aggregate information collected from several evaluations (Uusikylä & Virtanen, 2000; Vanhoof & Van Petegem, 2010), or for the quality control of evaluation studies (Hanssen, Lawrenz & Dunet, 2008; Lennie, Tacchi & Wilmore, 2012). According to Tsiatsos, Andreas and Pomportsis (2010), meta-evaluation is important in order to ensure that evaluations are valid, accurate and unbiased. Stufflebeam (2001:204) argues that "meta-evaluations help assure the integrity and credibility of evaluations". This view is supported by Wingate (2010), who also states that meta-evaluation aids in developing more sophisticated models and tools for evaluation.

Checklists exist for meta-evaluations in various subfields; some such instruments for judging program evaluations are:

- Program Evaluations Standards (Sanders & JCSEE, 1994), which present detailed criteria for determining utility, propriety, feasibility and accuracy of evaluations,
- Meta-Evaluation Checklist (Scriven, 2008), which proposes criteria for validity, credibility, propriety and cost-utility of evaluations, and
- American Evaluation Association Guiding Principles (Bamberger, Rugh & Mabry, 2006) that were developed for evaluators to guide the evaluation of a broader range of evaluations.

The aim of the meta-evaluation presented in this chapter was not to evaluate the findings of the evaluations performed by using the *DEVREF* Evaluation Framework, but to evaluate the evaluation framework itself. No standard meta-evaluation checklist was appropriate for evaluating *DEVREF,* due to its innovative and extensive nature. A custom-built meta-evaluation instrument was therefore developed by the researcher in order to evaluate both the *criteria* of the evaluation framework and the *methodology* applied for the evaluation. *DEVREF* was thus scrutinised by a meta-evaluation of its criteria and methods, with the aim of assessing its quality and suggesting possible improvements. The meta-evaluation was undertaken as a systematic review by separately evaluating the methodology employed and the criteria that comprise the framework.

9.2.2. The meta-evaluation questionnaire

A questionnaire was designed for this meta-evaluation and is attached as Appendix B-3. It contains five sections, of which the first four each have a number of evaluation statements to assess the framework in terms of its evaluation of instructional design principles, VR design principles, usability and context-specific design aspects, respectively. The fifth section deals with the effectiveness of the research method used for the evaluation of interactive VR e-training systems, namely, heuristic evaluation.

The meta-evaluation questionnaire was designed within the context of desktop VR training systems. The challenge was determining appropriate criteria for evaluating the merit of evaluation frameworks. The matter was addressed by selecting criteria from literature by acknowledged experts within the various terrains of instructional design, VR training, usability and context-specific aspects, as follows:

- Section 1 (Instructional Design): ten criteria relevant to desktop VR training systems were selected from Rogers *et al*. (2011), Mayer (2008), Zhang, Wang, Zhao, Li and Lou (2008) and Paas, Renkl and Sweller (2003). These criteria are feedback; visibility; constraints; consistency; affordance; contiguity; learner control; signalling; personalisation; and coherence.
- Section 2 (VR design): nine relevant criteria were selected from Stanney, Mollaghasemia, Reevesa, Breaux and Graeber (2003), Kalawsky (1999) and Tsiatsos, Andreas and Pomportsis (2010), i.e. interaction; navigation; object selection and manipulation; multimodal system output; visual output; auditory output; haptic output; presence; and engagement.
- Section 3 (Usability): ten relevant criteria were selected from Nielsen (1994), Squires and Preece (1999), and Shneiderman and Plaisant (2005), i.e. visibility of the system status; enable frequent users to use shortcuts; support internal locus of control; consistency and standards; error prevention; recognition rather than recall; aesthetic and minimalist design; design dialogues to yield closure; help users recognise, diagnose, and recover from errors; and help and documentation.
- Section 4 (Context-specific): four relevant criteria were selected from Alessi and Trollip (2001), Vrasidas (2004) and Nielsen (1994), i.e. learning in real-world contexts; corresponding concepts; appropriate language; and appropriate record keeping.

As previously stated, Sections 1 to 4 above evaluate the DEVREF evaluation criteria, whereas Section 5 moves beyond the criteria to evaluate the evaluation method employed. In the case in hand, the method was heuristic evaluation (HE). Ten evaluation statements were selected from advantages and disadvantages of heuristic evaluation mentioned in literature, and were taken from Ardito *et al.* (2006), Belkhiter, Boulet, Baffoun and Dupuis

(2003), Dix *et al.* (2004), the classic work of Jones, Scanlon, Tosunoglu, Morris, Ross, Butcher and Greenberg (1999), Karoulis and Pombortsis (2003), Kjeldskov, Skov and Stage (2004), and Preece (1993). The ten statements in Section 5 are:

1. HE is an effective evaluation method for identifying problems in the interaction design.
2. HE is an effective evaluation method that is relatively inexpensive to perform.
3. HE is an effective evaluation method that is relatively easy to perform.
4. HE can result in major improvements to a particular user interface.
5. During a short session, an expert evaluator can identify several usability problems.
6. A small number of experts can identify a range of usability problems.
7. Experienced evaluators can suggest solutions to usability problems that individual users may not pick up.
8. Expert evaluators may be biased due to their strong subjective views and preferences, and this may lead to biased reports.
9. It may be difficult to find evaluators who are experienced in both the specific domain of the system and HCI research.
10. Expert evaluation may not capture the variety of real users' behaviours. For example, novice users may perform unexpected actions that an evaluator might not think of.

According to Scriven (2008), a major requirement for determining the validity of an evaluation is whether the evaluation statements cover all the relevant considerations. Therefore, the statements in the meta-evaluation questionnaire were designed to evaluate the coverage of the selected criteria. The following format was used for Sections 1 to 4:

- The criterion was stated and briefly defined.
- The evaluators were requested to indicate whether the framework indeed evaluates this criterion, and also to substantiate their answers by indicating where this is done in the framework.
- At the end of each section the evaluators were requested to indicate whether any of the criteria in the meta-evaluation that were not addressed in the *DEVREF* Framework should indeed be included in the framework.
- The questionnaire also required the evaluators to indicate possible criteria that, according to their own expertise, should be included in the *DEVREF* Framework, but which were not addressed in the framework or the meta-evaluation questionnaire.

- Finally, space was provided for further comments the evaluators may have had relating *DEVREF*'s evaluation of the specific category of principles addressed by each section.

In Section 5, the evaluators were requested to indicate to what extent they agreed with each of ten statements relating to heuristic evaluation as a method for evaluating desktop virtual reality training systems. A five-point Likert scale, ranging from *Strongly agree* to *Strongly disagree* is used to indicate their answers.

The same six heuristic evaluators, who had used the *DEVREF* Framework to evaluate the two prototypes (discussed in Sections 8.4 and 8.5), were also requested to perform the meta-evaluation. Using these six experts ensured consistency in interpretation and, due to the selection of these evaluators, ensured that there were experts from all the areas covered by the meta-evaluation, including instructional design experts, VR training system developers, context-specific experts from the mining industry and usability experts. As indicated in Table 8.2 and explained in Section 8.4, the two usability experts (Evaluators A and B) both had previous experience in usability evaluation and instructional design, the mining training experts (Evaluators C and D) were experts in instructional design in the mining training environment, and the VR developers (Evaluators E and F) had extensive experience in such development. Furthermore, they all had the experience of applying *DEVREF* hands-on.

The meta-evaluation was conducted approximately six months after the heuristic evaluation of the second prototype, *ISGC*. The researcher visited each expert evaluator individually, explained the purpose and procedure of the meta-evaluation, and supplied them with a copy of the *DEVREF* Framework for reference purposes during their meta-evaluation. Each expert then completed the meta-evaluation questionnaire using as much time as they required.

9.3. Findings of the meta-evaluation

The meta-evaluation was applied as formative evaluation, with the aim of possible improvement of *DEVREF*. The next two subsections discuss the findings of the meta-evaluation with regard to the framework's criteria and methodology respectively.

9.3.1. Findings of the meta-evaluation of the *DEVREF* criteria

The findings are presented per section, in the same sequence as they are structured in the meta-evaluation questionnaire.

Tables 9.1 to 9.4 summarise the responses of the experts to each of the criteria, one table per section. The six evaluators are indicated as Evaluator A through F, in the same order as was reported in the findings of the heuristic evaluation of the prototypes in Chapter Eight. The same colour-coded scheme is used as in the presentation of the findings of the heuristic evaluations of *LSF* and *ISGC* in Chapter Eight, that is, the usability experts' responses are in red, the mining experts in green and the VR experts in blue.

Where an evaluator agreed that the framework did assess the criterion being evaluated, he indicated which evaluation statements in the framework addressed that particular criterion, for example, in Table 9.1, Evaluator A indicated that *DEVREF* assessed *Feedback* in Section 1.3 and in Section 2.2, specifically via Evaluation Statement 2.2.1. Where no entries are made in a specific column, it indicates that the evaluator could not find an evaluation statement in *DEVREF* that evaluated the criterion. In further discussions, 'Evaluation Statement' is abbreviated as ES.

9.3.1.1. Instructional design criteria

Table 9.1 shows that all the evaluators consistently determined that all the instructional design criteria in the meta-evaluation questionnaire were assessed by *DEVREF*, except for *Affordance*. Only evaluators A and E felt that ES 2.1.3 assessed affordance and the other four evaluators did not agree.

Table 9.1: Meta-evaluation results of Section 1 – Instructional Design.

Criteria	Evaluators' feedback					
	A	B	C	D	E	F
Feedback	**1.3, 2.2.1**	**1.3**	**1.3**	**1.3**	**1.3**	**1.3**
Visibility	**2.3.3, 2.6.3**	**2.6.3**	**2.6.3, 2.3.3**	**2.6.3, 2.3.3**	**2.3.3, 2.6.2, 2.6.3**	**2.3.3, 2.6.2, 2.6.3**
Constraints	**2.7.1, 2.2.1, 2.2.2, 2.2.3**	**2.2.1, 2.2.2, 2.2.3**	**2.7.1, 2.2.1**	**2.7.1, 2.2.1, 2.2.2**	**2.2.1, 2.7.1**	**2.2.1, 2.7.1**

Consistency	**2.3**	**2.3**	**2.3**	**2.3**	**2.3**	**2.3**
Affordance	**2.1.3**				**2.1.3**	
Contiguity	**1.6.1, 1.8.3**	**1.6.1, 1.8.3**	**1.8.3**	**1.8.3**	**1.8.3**	**1.6.1, 1.8.3**
Learner Control	**1.8.2, 3.1.1, 3.1.2**	**1.8.2, 3.1.1, 3.1.2**	**1.8.2**	**1.8.2, 3.1.1**	**1.8.2, 3.1.1**	**1.8.2, 3.1.1**
Signalling	**1.6.1**	**1.6.1**	**1.6.1**	**1.6.1**	**1.6.1**	**1.6.1**
Personalisation	**1.7.4**	**1.7.4**	**1.7.4**	**1.7.4**	**1.7.4**	**1.7.4**
Coherence	**1.6.2**	**1.6.2**	**1.6.2**	**1.6.2**	**1.6.2**	**1.6.2**

There was close consensus in the heuristics that the experts selected as criteria that evaluated the various facets of instructional design. However, in a few cases an evaluator perceived additional richness in a criterion that evaluated some other aspect, for example, Evaluator A felt that ES 2.2.1 also related to *Feedback* and not only ES 1.3, which was named by all six evaluators.

At the end of each section there were three open-ended questions:

1. If any of the above criteria are not evaluated by the *DEVREF* Framework, are you of the opinion that those criteria should be included in the framework?
2. Do you think any other instructional design criteria should be added to the *DEVREF* Framework? If so, please indicate which criteria should be added.
3. Do you have any other comments relating to the *DEVREF* Framework's evaluation of instructional design principles for desktop VR training systems?

In answering Question 1, all the evaluators who did not find an evaluation statement that assessed *Affordance* indicated that this should indeed be addressed in the framework. Evaluator B also commented regarding the *Feedback* criterion, that the framework did not explicitly evaluate cognitive feedback. Regarding Question 2, none of the evaluators proposed new criteria. The only feedback on Question 3 was from Evaluator C, who queried why some answers to Section 1, which covers instructional design criteria, were found in other categories in the framework, namely Category 2 (General usability) and Category 3 (VR design).

Interpretation of findings

In addressing certain factors, some evaluators found more supporting evaluation statements than others, for example, in evaluating *Constraints,* Evaluators C, E and F indicated that this was assessed by ES 2.7.1 and ES 2.2.1, but Evaluator A also indicated 2.2.2 and 2.2.3 as relevant to assessing constraints, and Evaluator B indicated 2.2.1, 2.2.2 and 2.2.3 only.

Section 9.2.2 explained that the meta-evaluation questionnaire was designed to evaluate the coverage of the selected criteria, therefore the number of evaluation statements included under a criterion is less important than the fact that the criterion is indeed assessed. In the framework, different aspects of a criterion can be evaluated, resulting in a word occurring more than once. However, where one criterion is assessed multiple times, it could indicate duplication in the framework, that is, multiple entries in Table 9.1 could highlight a need to refine the framework.

To investigate the difference of opinion between the evaluators regarding the *Affordance* criterion, the researcher referred back to ES 2.1.3 mentioned by two of the evaluators, as well as to the original intention of this statement when the framework was developed. ES 2.1.3 is part of the subsection *Functionality* and was phrased as 'icons, labels and symbols are intuitive and meaningful to trainees, bearing in mind the level of trainee context and experience'. In the meta-evaluation questionnaire, *Affordance* is defined as 'the aspect in the design of an object that suggests how the object should be used. Objects and systems should be designed with attributes that support users in knowing how to use them, e.g. scroll bars afford moving up and down'. The evaluators who indicated 2.1.3 as relevant to assessing affordance, both pointed out that the use of the word 'intuitive' caused them to believe that it could be related to affordance. However, Evaluator C also indicated that 2.1.3 'did not address affordance sufficiently'. The above implies that, in the interest of clarity, adding affordance to the framework as a separate evaluation statement would be an improvement to the *DEVREF* Framework.

Regarding Evaluator B's comment that 'cognitive feedback' was not explicitly evaluated, the framework evaluates feedback in several different evaluation statements in Category 1 (see Appendix B-1):

- 1.3.1: constructive and supportive feedback on performance

- 1.3.2: relevant feedback
- 1.3.3: level of achievement
- 1.3.4: feedback on incorrect responses.

Furthermore, in Categories 2 and 3, other evaluation statements are found that relate to feedback:

- 2.4.1: error messages
- 2.4.2: help to recover from cognitive errors
- 2.4.3: clear indications of errors and constructive instructions for recovery
- 3.2: multimodal system feedback.

These statements from Categories 1, 2 and 3 should be sufficient to address the apparent lack of evaluating cognitive feedback.

Regarding the comment that there are related evaluation statements in different categories of the framework, it should be noted that aspects of some criteria are relevant to different categories. The complete framework is used when evaluating a system, and repeating a criterion would lead to redundancy. The categories were not designed to be mutually exclusive, as certain criteria may be relevant to different categories, but will only be listed once in the framework.

Consolidation of these findings and associated recommendations is given in Section 9.3.3, which summarises the findings of the meta-evaluation questionnaire.

9.3.1.2. Virtual reality design criteria

Table 9.2 summarises the results of the VR design category. The experts' evaluations show that all the criteria were sufficiently evaluated by the framework, except for *Haptic output*. In answering the first of the open-ended questions at the end of the section, two evaluators indicated that haptic output is not relevant to desktop VR and need not be incorporated in the framework. The others suggested that, due to current research in this area, it may be part of future developments and that the framework should make provision for it.

Table 9.2: Meta-evaluation results of Section 2 – Virtual Reality Design.

Criteria	Evaluators' feedback					
	A	B	C	D	E	F
Interaction	**1.6.2, 2.8**	**2.8**	**1.4.5, 2.8**	**1.4.5, 2.8**	**1.4.5, 2.8**	**1.4.5, 2.8**
Navigation	**3.5**	**3.5**	**3.5**	**3.5**	**3.5**	**3.5**
Object selection and manipulation	**3.6**	**3.6**	**3.6**	**3.6**	**3.6**	**3.6**
Multimodal system output	**3.2**	**3.2**	**3.2**	**3.2**	**3.2**	**3.2**
Visual output	**3.2.1, 3.2.3**	**3.2.1, 3.2.3**	**3.2.1, 3.2.3**	**3.2.1, 3.2.3**	**3.2.1, 3.2.3**	**3.2.1, 3.2.3**
Auditory output	**3.2.4, 3.2.5**	**3.2.4, 3.2.5**	**3.2.4, 3.2.5**	**3.2.4, 3.2.5**	**3.2.4, 3.2.5**	**3.2.4, 3.2.5**
Haptic output						
Presence	**3.3**	**3.3**	**3.3**	**3.3**	**3.3**	**3.3**
Engagement	**1.4.4, 1.4.5, 3.7.3, 3.4.1**	**1.4.4, 1.4.5**	**1.4.4, 1.4.5, 3.4.1**	**1.4.4, 1.4.5**	**1.4.4, 1.4.5**	**1.4.4, 1.4.5, 3.4.1**

From Table 9.2, it is clear that the evaluation of this category was consistent in that the responses were almost identical across the spectrum, except for *Engagement*, where Evaluators A, C and F also listed ES 3.4.1 as relevant, while Evaluator A also noted ES 3.7.3. There was a marginal difference in the *Interaction* criterion, where Evaluator A also mentioned ES 1.6.2, but neither Evaluator A nor Evaluator B included 1.4.5.

In answering the open-ended Question 2 about proposing new VR design criteria for *DEVREF*, the only feedback came from the evaluators who have VR system design experience. Two evaluators proposed adding a criterion to evaluate the inclusion of a captivating storyline. Another proposal was to evaluate the use of logical barriers in areas where physical barriers were absent, but to which trainees should not have access, for example, having an obstruction in the path of an open doorway to prevent entry. The only comment for Question 3 was Evaluator C commenting that Evaluation Statement 3.4.1 was written in a different style from the others.

Interpretation of findings

The comments from the evaluators regarding haptic output warranted further investigation into the introduction of this for future desktop VR systems. Haptic output refers to the ability of a system to provide a sense of touch or physical sensation.

The researcher identified some haptic technologies that could be appropriate to desktop VR systems (Robles-De-La-Torre, 2009; Geomagic.com, 2013; Evangelho, 2014). In the literature, the terminology *tactile feedback* or *force feedback* is also used to describe haptic output. Some haptic devices used with desktop applications are force feedback joysticks, steering wheels and mice. Such devices are popular within the computer gaming industry, as they provide the gamer with a sense of realism. An example of haptic output is a steering wheel that starts to vibrate if the vehicle being 'driven' during a simulation, leaves the road. More advanced haptic systems can provide pneumatic feedback, by using air jets to create a wind effect, or electrotactile feedback, where electrodes are attached to the user's fingers to provide electrical pulses (Berkley, 2003; Stanney *et al.*, 2003).

The inclusion of haptic technologies in a desktop VR system entails using additional haptic devices other than the standard monitor screen or computer mouse to supply the haptic output. It is indeed possible to develop desktop VR systems that provide haptic output using these haptic devices. It is therefore concluded that adding haptic output to *DEVREF* would be relevant and an improvement. Haptic output can be accommodated as an additional evaluation statement to the existing *Multimodal system output/feedback* heuristic, which would then cover audio, visual and haptic output.

Regarding the proposal to add a criterion to assess the storyline, the researcher is of the opinion that this would impact the generality and transferability of the framework, since not all VR training systems require a script or are story-based. Some virtual environments could be used for orientation purposes or trainees could cognitively construct knowledge for themselves by interacting with the virtual environment and observing the consequences of their actions. It was therefore decided not to add a new criterion to the framework, but to cater for storyline assessment by adding an appropriate evaluation statement to an existing criterion in the framework.

In reviewing the request by Evaluator D for the framework to evaluate the presence of logical barriers, it became clear that this is a navigational issue not specifically evaluated by the framework. Since *DEVREF* already contains a *Navigation* criterion, it can be solved by adding an evaluation statement to this effect to the *Navigation* criterion.

Evaluator C's comment on the style of ES 3.4.1 is valid. The statement 'users do not find it difficult to maintain knowledge of their location' differs from all the other evaluation statements in that it is written in a negative format, while the others are presented in a positive format, for example, 'input devices are easy to use and control' and 'it is clear to the user how to exit'. ES 3.4.1 should therefore be modified to read 'users find it easy to maintain knowledge of their location'.

All the findings and recommendations are consolidated in Section 9.3.3.

9.3.1.3. System usability criteria

In evaluating the system usability criteria, the expert evaluators consistently gave the same or very similar responses, which are presented in Table 9.3. They all agreed that two criteria were not assessed by the *DEVREF* Framework, *Shortcuts for frequent users,* and *Design dialogues to yield closure*. In the open-ended section, all the evaluators pointed out that *Shortcuts for frequent users* should be added to the framework, but there were differences in opinion regarding *Design dialogues to yield closure*. Evaluator A felt the matter was sufficiently addressed in the evaluation statements relating to feedback; Evaluator B stated that 'this definition is vague and confusing and there is no need for it'; Evaluator C believed that the criterion should be added, while the other evaluators did not think it was necessary within desktop VR systems.

The only new criterion was proposed by Evaluator E, who suggested that provision should be made for evaluating whether the system permits jumping from one location to another within a large area map. There were no other comments.

Table 9.3: Meta-evaluation results of Section 3 – System Usability.

Criteria	Evaluators' feedback					
	A	**B**	**C**	**D**	**E**	**F**
Visibility of system status	**2.5**	**2.5**	**2.5.1**	**2.5**	**2.5**	**2.5**
Shortcuts for frequent users						
Support internal locus of control	**3.1**	**3.1**	**3.1**	**3.1**	**3.1**	**3.1**
Consistency and standards	**2.3, 2.1.3**	**2.3**	**2.3**	**2.3**	**2.3**	**2.3**
Error prevention	**2.7**	**2.7**	**2.7**	**2.7**	**2.7**	**2.7**
Recognition rather than recall	**2.3.3**	**2.3.3**	**2.3.3**	**2.3.3**	**2.3.3**	**2.3.3**
Aesthetic and minimalist design	**2.6.4, 1.6.2**	**2.6, 1.6.2**	**2.6**	**2.6**	**2.6.1, 2.6.4**	**2.6**
Design dialogues to yield closure						
Recognise, diagnose, and recover from errors	**2.4.2, 2.4.3**	**1.3, 2.4, 2.5**	**2.4.2, 2.4.3**	**2.4.2, 2.4.3**	**2.4.2, 2.4.3**	**2.4.2, 2.4.3**
Help and documentation	**2.2.2, 2.2.4, 4.2**	**2.2.2, 4.2**	**2.2.2, 4.2**	**2.2.2, 4.2**	**2.2.2, 2.2.4**	**2.2.2, 2.2.4, 4.2**

Interpretation of findings

Since all the evaluators agreed that *Shortcuts for frequent users* should indeed be assessed by the framework, this criterion will be incorporated to improve *DEVREF*. Regarding the uncertainty relating to *Design dialogues to yield closure,* only one evaluator deemed it necessary to add it while the other five considered it unnecessary or redundant due to it being adequately addressed through other evaluation statements. The argument in support of including it, is therefore not strong enough and it will not be added.

Regarding one evaluator's suggestion for easy relocation within large area maps, it is clear that this would be beneficial to applications using maps or systems that make provision for instant position relocation by user control. For example, a user could instantly move from one location in the mine to another by indicating a desired position on a map. As this would apply only to some applications and is not a generic feature in all desktop VR training systems, this aspect will be covered by adding an additional evaluation statement to a relevant existing criterion, rather than including a new criterion. For consolidation and summary of all the findings, please see Section 9.3.3.

9.3.1.4. Context-specific design criteria

Table 9.4 presents results of the meta-evaluation of context-specific design criteria. In this section there were only marginal differences between the evaluators' responses, and no comments emerged from the open-ended questions.

Table 9.4: Meta-evaluation results of Section 4 – Context-specific Design.

Criteria	Evaluators' feedback					
	A	B	C	D	E	F
Learning in real-world contexts	**4.1**	**4.1**	**4.1**	**4.1**	**4.1**	**4.1**
Corresponding concepts	**4.7.4, 4.8**	**4.7.4**	**4.7.4, 4.8**	**4.7.4, 4.8.2**	**4.7.4, 4.8.2, 4.8.3**	**4.7.4, 4.8.2, 4.8.3**
Appropriate language	**4.7.5, 4.8.2**	**4.7.5, 4.8.1, 4.8.2**	**4.7.5**	**4.7.5**	**4.7.5**	**4.7.5**
Appropriate record keeping	**4.5**	**4.5**	**4.5**	**4.5**	**4.5**	**4.5**

Interpretation of findings

The findings were almost entirely consistent across evaluators, indicating that all the aspects in this section were assessed by the *DEVREF* Framework. With no additional criteria proposed and no further comments received, it can be concluded that this section is complete and acceptable.

9.3.2. Findings of the meta-evaluation of the *DEVREF* method

In Section 5 of the meta-evaluation, the expert evaluators were requested to indicate to what extent they agreed with the application of the specific evaluation method, in this case Heuristic Evaluation (HE), for evaluating desktop virtual reality training systems. The results are given in Table 9.5. The table contains the ten statements from the meta-evaluation and indicates, in the standard colour code, which option each evaluator selected.

Table 9.5: Meta-evaluation results of Section 5 – HE as a suitable evaluation method.

1. HE is an effective evaluation method for identifying problems in the interaction design.

Strongly Agree	Agree	Maybe	Disagree	Strongly Disagree
B, D	**A, C, E, F**			

b. HE is an effective evaluation method that is relatively inexpensive to perform.

Strongly Agree	Agree	Maybe	Disagree	Strongly Disagree
A, C	**D, E, F**	**B**		

c. HE is an effective evaluation method that is relatively easy to perform.

Strongly Agree	Agree	Maybe	Disagree	Strongly Disagree
C, D, F	**A, B, E**			

d. HE can result in major improvements to a particular user interface.

Strongly Agree	Agree	Maybe	Disagree	Strongly Disagree
C	**A, D, F**	**B, E**		

e. During a short session, an expert evaluator can identify several usability problems.

Strongly Agree	Agree	Maybe	Disagree	Strongly Disagree
	A, B, C, D, F	**E**		

f. A small number of experts can identify a range of usability problems.

Strongly Agree	Agree	Maybe	Disagree	Strongly Disagree
A, C	**B, D, E, F**			

g. Experienced evaluators can suggest solutions to usability problems that individual users may not pick up.

Strongly Agree	Agree	Maybe	Disagree	Strongly Disagree
A, **B**, **C**, **E**	**D**, **F**			

h. Expert evaluators may be biased due to their strong subjective views and preferences, and this may lead to biased reports.

Strongly Agree	Agree	Maybe	Disagree	Strongly Disagree
	C	**A**, **D**, **E**, **F**		**B**

i. It may be difficult to find evaluators who are experienced in both the specific domain of the system and HCI research.

Strongly Agree	Agree	Maybe	Disagree	Strongly Disagree
A, **B**	**C**, **D**, **E**, **F**			

j. Expert evaluation may not capture the variety of real users' behaviours. For example, novice users may perform unexpected actions that an evaluator might not think of.

Strongly Agree	Agree	Maybe	Disagree	Strongly Disagree
A	**B**, **C**, **D**, **E**, **F**			

From Table 9.5 it is clear that in most cases the experts were closely in agreement, except for Statement 8: 'Expert evaluators may be biased due to their strong subjective views and preferences, and this may lead to biased reports'. Four experts indicated that this may be possible and Evaluator C agreed with the statement, but Evaluator B strongly disagreed that heuristic evaluators could be biased.

The evaluators were also requested to add unprompted comments relating to the use of the heuristic evaluation method for evaluating desktop VR training systems. Only Evaluator A commented, suggesting that the ideal scenario would be to have heuristic evaluators with knowledge in all the relevant areas of the framework, that is, experts proficient in instructional design, usability, VR system design and the context wherein systems will be evaluated using the framework. However, it is rare to obtain a so-called 'double expert' and would be even harder to acquire expert evaluators with expertise across the spectrum.

Interpretation of findings

Regarding possible bias from heuristic evaluators, it is evident from literature that other researchers have found that evaluators could be biased (Metzger, Flanagin & Medders, 2010; Nielsen, 1994; Pinelle, Wong & Stach, 2008; Preece, 1993).

The responses in Table 9.5 show that, in general, the evaluators were in agreement with the statements relating to the use of the heuristic evaluation research method. It can therefore be determined that this approach is suitable for evaluating desktop VR training systems.

9.3.3. Consolidation of findings of the meta-evaluation questionnaire

The findings in Section 9.3.1 of the meta-evaluation criteria point towards improvements that can be made to the *DEVREF* Framework. In summary, the following revisions should be implemented:

- From the findings of the instructional design section, it was concluded that *affordance* should be added to the framework as a separate evaluation statement in a relevant criterion. Since two of the evaluators mentioned ES 2.1.3, which belongs to the *Functionality* criterion, as relevant to this topic, a new evaluation statement on affordance can be added to this criterion.
- The evaluators' feedback on virtual reality design indicated some potential improvements:
  - *Haptic output* should be included as a new evaluation statement in the *Multimodal system output/feedback* criterion.
  - Evaluation of a *captivating storyline* should be an additional evaluation statement in a relevant existing criterion. Since a captivating storyline relates to motivation and increased trainee participation, it was decided to add the statement to the *Motivation and Creativity* criterion.
  - The presence of *logical barriers* should be a new evaluation statement in the *Navigation* criterion.
  - ES 3.4.1 of the *Orientation* criterion should be rephrased in a positive format similar to the other evaluation statements.
- The system usability section indicated that two modifications were required:
  - Providing *shortcuts to frequent users* refers to experienced users who desire to reduce the number of interactions and increase the speed of

accomplishing tasks. This issue can be accommodated by adding an evaluation statement to the *Various user modes* criterion.

- Provision for *map-based relocation* should be made by adding such an evaluation statement to a relevant criterion. Since this relates to moving and repositioning in the virtual environment, it was decided to do this within the *Navigation* criterion.

The findings of the evaluation of the context-specific section did not indicate any improvements to the framework.

As was mentioned in Section 9.3.1, the issue that emerged during the meta-evaluation of more than one evaluation statement assessing a criterion, could be due to different aspects of a criterion being evaluated by the framework, or might point towards duplication in the framework. The researcher investigated multiple entries and is satisfied that in each case a different aspect is evaluated and that multiple findings were not due to duplication in the framework. The next section reports on the implementation of the improvements discussed in this section.

9.4. The *DEVREF* framework revisited

This section addresses Research Subquestion 5 of this study, along with Sections 3.6.3 and 5.8:

RQ5	What structure, categories and criteria should be incorporated in an evaluation framework for virtual reality training systems in the mining industry?

The findings of the meta-evaluation, as discussed in Sections 9.3.1 and 9.3.2, clearly indicate that the meta-evaluation exercise was meaningful and that it achieved its aim in determining whether the evaluation statements in the *DEVREF* Framework cover all the relevant considerations. Using the feedback received, the researcher was able to improve the *DEVREF* Framework. An additional indirect outcome, which was not anticipated prior to the meta-evaluation, is that the evaluators' comments also focused attention on the wording of criteria. Ambiguities were resolved and more appropriate terms were suggested where necessary. This should result in easier use of *DEVREF*.

In Chapter Eight, the findings of the heuristic evaluations of the *LSF* and *ISGC* prototypes were presented and discussed. The interpretation of these findings, as well as the findings of the meta-evaluation described in this chapter, all revealed necessary improvements to the *DEVREF* Evaluation Framework.

Some of the improvements suggested in this chapter involve only modifications to the heuristic evaluation instrument based on the *DEVREF* Framework, and not the framework itself. However, any modifications to the actual *DEVREF* Framework also lead to modifications to the associated heuristic evaluation instrument.

Section 9.3.3 outlined the revisions required in the framework as a result of the meta-evaluation. Chapter 8 reports on the application of *DEVREF* to evaluate the prototypes. These evaluations also revealed required improvements to the DEVREF Framework and the accompanying heuristic evaluation instrument:

- In the analysis of findings of the heuristic evaluation of *LSF* in Section 8.4.1, two evaluators commented that the word 'performance' in ES 2.2 of the instructional design category could also refer to 'knowledge'. It was therefore decided to use both in the revised evaluation statement, resulting in: 'The assessment opportunities will serve to enhance trainees' performance and knowledge'.
- From the findings of the heuristic evaluation of *ISGC* in Section 8.5.1, the following improvements were recommended:
  - ES 5.2 in the instructional design category should be rephrased to prevent misunderstanding of whether the statement refers to level of content or level of users. The following rephrased statement should clarify the issue: 'In terms of content, the system caters for novice and knowledgeable trainees'.
  - ES 7.3, also from the same category, should be clarified to define what is meant by 'sufficient scaffolding'. To resolve the issue, it was decided to omit the word sufficient' from the statement, to read 'Scaffolding support is provided (in the form of hints, prompts and feedback) to help trainees achieve training goals'.
  - ES 2.2 in the general usability category caused confusion in that it evaluated both accessibility and appropriateness of user help in one statement. This can be resolved by dividing this evaluation statement into two separate statements. ES 2.2 will be rephrased as 'Help for operating the program is

accessible at any time' and a new ES 2.5 will be added at the end of the section to state: 'Help for operating the program is appropriate'.

- It was indicated that ES 8.1 in the VR system design category duplicated ES 5.2 in the instructional design category. It was decided to delete ES 8.1: 'The system employs various user modes to cater for a range of users, from novices to experts'.

In addition, minor syntactic and semantic weaknesses were corrected by the researcher or a critical reader.

The above-mentioned improvements, as well as those mentioned in Section 9.3.3, were incorporated in the *DEVREF* Framework and the accompanying heuristic evaluation instrument. The improved *DEVREF* Framework is shown in Table 9.6. The revised heuristic evaluation instrument is attached as Appendix B-4. In Table 9.6, the references in blue relate to sources published after the inception of the original framework, but they are in harmony with the original references and support and strengthen the criteria developed for *DEVREF.*

Table 9.6: Improved Heuristic Evaluation Framework for Desktop VR Training Applications.

Category 1: Instructional Design		
	Heuristic	**Literature References**
1	**Clear goals, objectives or outcomes:** • The training program makes it clear to the learner what is to be accomplished and what will be gained from its use. • There are clear goals, objectives or outcomes for the training program. • Clear goals, objectives or outcomes are communicated at the beginning of the training program. • The outcomes are measurable.	Ritchie and Hoffman (1997), Albion (1999), Wein, Piccirilli, Coffey and Flemming (2000), Alessi and Trollip (2001), Reeves, Benson, Elliot, Grant, Holschuh, Kim, Kim, Lauber and Loh (2002), McLoughlin, in Edmundson (2003), Ardito, Costabile, De Marsico, Lanzilotti, Levialdi, Plantamura, Roselli, Rossano and Tersigni (2004b).

2	**Instructional assessment:** • The program provides assessment opportunities that are aligned with the objectives or outcomes. • The assessment opportunities will serve to enhance trainees' performance and knowledge.	Albion (1999), Patel, Stefani, Sharples, Hoffmann, Karaseitanidi and Amditis (2006).
3	**Feedback to user responses:** • The training program provides trainees with constructive and supportive feedback on their performance • The feedback is relevant to the training content. • The feedback informs the trainee regarding his level of achievement in the training program. • The feedback indicates incorrect responses and provides information on the correct responses.	Alessi and Trollip (2001), Vrasidas (2004), Rusu, Muñoz, Roncagliolo, Rudloff, Rusu and Figueroa (2011), Munoz & Chalegre (2012).
4	**Motivation and creativity**: • The system supports intrinsic motivation by providing challenges to trainees • The system provides encouragement when errors are made • The program captures the trainee's attention early and retains it throughout. • This training program increases trainees' confidence by providing them with reasonable opportunities to accomplish the objectives successfully. • The program engages trainees by its relevant content. • The program engages trainees by its interactivity. • The program has a captivating storyline.	Albion (1999), Alessi and Trollip (2001), Reeves *et al.* (2002), Chalmers (2003), Vrasidas (2004), De Villiers (2005a), Ssemugabi and De Villiers (2007), Magner, Schwonke, Aleven, Popescu and Renkl (2013), Mayer (2014).
5	**Differences between individual users:** • The system takes account of linguistic and cultural differences by allowing trainees to select between different languages. • In terms of content, the system caters for both novice and knowledgeable trainees. • The system caters for trainees with different levels of computer experience.	Alessi and Trollip (2001), Barber (2002), Reeves *et al.* (2002), Chalmers (2003), Liu, in Edmundson (2003), McLoughlin, in Edmundson (2003), Kamppuri, Tedre and Tukiainen (2006), Rogers *et al.* (2011), Lau, Yen, Li & Wah (2014).
6	**Reduction of extraneous processing in working memory:** • The training program effectively uses signalling to highlight essential issues (e.g. restating important points, using headings for important points, or stressing them in audio mode). • Redundancy is avoided i.e. unnecessary information is not presented. • Redundancy and overload are avoided by not	Chalmers (2003), Mayer (2008), Hollender, Hofmann, Deneke and Schmitz (2010), Sweller, Ayres and Kaluga (2011), Morrison, Dorn and Guzdial (2014), Mason, Cooper and Wilks (2015).

	reiterating the same material in multiple modes (e.g. the program presents information using pictures and spoken words, rather than presenting it in pictures, spoken words, and printed words).	
7	**Fostering of germane cognitive load (germane cognitive load is the load devoted to the processing, construction and automation of schemas):** • The training program supports the formation of mental schema by explaining where newly acquired knowledge fits into the bigger picture. • The system encourages encoding of the training content in long-term memory by presenting questions after each learning segment. • Scaffolding support is provided (in the form of hints, prompts and feedback) to help trainees achieve training goals. • The training program presents narration in a colloquial conversational style. • The training program prompts trainees to link concrete example information for each problem category to more abstract information.	Albion (1999), Sweller (1999), Alessi and Trollip (2001), Chalmers (2003), Ardito *et al.* (2004b), Van Merriënboer and Sweller (2005), Sawicka (2008), Bennett, Stothard and Kehoe (2010), Hollender *et al.* (2010), Sweller, Ayres and Kaluga (2011), Teräs and Herrington (2014).
8	**Appropriate intrinsic cognitive load**: • Working through the training program does not cause trainees to split their attention between multiple sources of visual information. • The program enhances retention by presenting information in learner-paced segments, rather than as a continuous presentation. • The system effectively supports dual channel processing of simultaneous visual and verbal material.	Alessi and Trollip (2001), Pollock, Chandler and Sweller (2002), Reeves *et al.* (2002), Shneiderman and Plaisant (2005), Mayer (2008), Zhang, Wang, Zhao, Li and Lou (2008), Sweller, Ayres and Kaluga (2011), Munoz & Chalegre (2012), Lau, Yen, Li & Wah (2014), Morrison, Dorn and Guzdial (2014).
Category 2: General Usability		
	Heuristic	**References**
1	**Functionality**: • The interface provides the level of functionality the user requires to complete a task. • The interface provides adequate back button functionality to return to a previous screen. • Icons, labels and symbols are intuitive and meaningful to trainees, bearing in mind the level of trainee context and experience. • Objects are designed with attributes that support affordance.	Kalawsky (1999), Dringus and Cohen (2005), Adebesin, Kotze and Gelderblom (2010), Hvannberg, Halldórsdóttir and Rudinsky (2012).

2	**User guidance:** • The interface provides clear indications of what the next required action will be. • Help for operating the program is accessible at any time. • Trainees receive clear instructions on how to use the training program. • Guidance to solve problems is given in the form of examples, diagrams, videos or photographs. • Help for operating the program is appropriate.	Nielsen (1994), Kalawsky (1999), Alessi and Trollip (2001), Reeves *et al.* (2002), Dringus and Cohen (2005), Adebesin *et al.* (2010), Guimarães & Martins (2014), Lau, Yen, Li & Wah (2014).
3	**Consistency:** • There is consistency in the sequence of actions taken in similar situations. • There is consistency in the use of images, prompts, screens, menus, colours, fonts and layouts. • Objects, options, and permissible actions are visible so that users do not have to remember instructions. • Different screens that have similar operations, use similar elements for achieving similar tasks.	Nielsen (1994), Kalawsky (1999), Squires and Preece (1999), Reeves *et al.* (2002), Dix *et al.* (2004), Dringus and Cohen (2005), Shneiderman and Plaisant (2005), Wong, Marcus, Ayres, Smith, Cooper and Paas (2009), Adebesin *et al.* (2010), Olsen (2010), Hvannberg *et al.* (2012), Munoz & Chalegre (2012), Guimarães & Martins (2014).
4	**Error correction:** • Error messages are expressed in plain language. • Learners are provided with the necessary help to recover from cognitive errors. • Error messages indicate precisely what the problem is and give simple, constructive, specific instructions for recovery.	Nielsen (1994), Kalawsky (1999), Squires and Preece (1999), Powell (2001), Reeves *et al.* (2002), Karoulis and Pombortsis (2003), Dix *et al.* (2004), Shneiderman and Plaisant (2005), Adebesin *et al.* (2010), Rusu *et al.* (2011), Guimarães & Martins (2014).
5	**System status**: • The training program keeps the trainee informed about what is going on through constructive, appropriate and timely feedback. • For every action taken by the trainee, there is a visual or audio response by the training program so that learners can see and understand the results of their actions. • The program responds to actions initiated by the user and there are no surprise actions from the system's side.	Nielsen (1994), Levi and Conrad (1996), Albion (1999), Squires and Preece (1999), Reeves *et al.* (2002), Dix *et al.* (2004), Shneiderman and Plaisant (2005), Rusu *et al.* (2011), Guimarães & Martins (2014).

6	**Error prevention:** • The training program is designed in such a way that the learner cannot easily make serious errors. • When the learner makes an error, the system responds with an error message. • Trainees can recognise situations where errors occur due to the way they provided input, and not due to incorrect content in their response. • The system is robust and reliable throughout.	Nielsen (1994), Kalawsky (1999), Squires and Preece (1999), Alessi and Trollip (2001), Powell (2001), Reeves *et al.* (2002), Karoulis and Pombortsis (2003), Dix *et al.* (2004), Dringus and Cohen (2005), Shneiderman and Plaisant (2005), Rusu *et al.* (2011), Munoz & Chalegre (2012).
7	**Aesthetics:** • The screens are pleasing to look at. • The buttons and selections are of a size that is adequately viewable. • The text is of a size that is adequately viewable. • There is not too much content or information on the screens.	Alessi and Trollip (2001), Reeves *et al.* (2002), Dringus and Cohen (2005), Magner *et al*. (2013), Guimarães & Martins (2014). .
8	**Interactivity:** • The training program uses clear and simple terminology that supports trainees in understanding how to interact with the system. • The program provides interactions that support trainees in learning the necessary content. • Working through the program requires regular trainee interactivity to maintain attention and facilitate comprehension.	Alessi and Trollip (2001), Preece, Rogers and Sharp (2002), Dringus and Cohen (2005), Adebesin *et al.* (2010), Su, Wang, Wu & Kuo (2013), Schofield (2014), Lau, Yen, Li & Wah (2014), Mason, Cooper and Wilks (2015).
Category 3: Virtual Reality System Design		
	Heuristic	**References**
1	**User control:** • The user is able to interact with, or control, the virtual environment in a natural manner. • Responses from the environment to the participant's control actions and movements, are perceived as immediate or close-to-immediate. • The system permits easy reversal of actions. • Trainees are able to exit the system at any time when they need to do so.	Nielsen (1994), Kalawsky (1999), Squires and Preece (1999), Dix *et al.* (2004), Shneiderman and Plaisant (2005), Wilson and D'Cruz (2006), Schofield and Lester (2010), Rebelo & Noriega (2012), Guimarães & Martins (2014), Lau, Yen, Li & Wah (2014).

2	**Multimodal system output/feedback:** • The effect of the trainee's actions on objects in the virtual environment, is immediately visible and conforms to the laws of physics and the trainee's perceptual expectations. • The visual representation of the virtual world maps to the trainee's normal perception of that environment. • Distortions are not noticeable in visual images • Audio is integrated seamlessly into user task activity. • Audio information is meaningful and timely. • The system provides appropriate haptic output.	Mereu and Kazman (1996), Oshhima, Yamamoto and Tamura (1996), Richard, Birebent, Coiffet, Burdea, Gomex and Langrana (1996), Kalawsky (1999), Hix and Gabbard (2002), Sutcliffe and Gault (2004), Hanson & Shelton (2008), Lau, Yen, Li & Wah (2014), Schofield (2014).
3	**Presence:** • Users feel as if they are part of the virtual environment and not isolated from it. • The virtual environment experience is consistent with similar real-world experiences.	Witmer and Singer (1998), Kalawsky (1999), Sadowski and Stanney (2002), Bowman and McMahan, (2007), Hanson & Shelton (2008), Rebelo & Noriega (2012), Su, Wang, Wu & Kuo (2013).
4	**Orientation:** • Users find it easy to maintain knowledge (or 'awareness') of their location while moving through the virtual environment. • The virtual environment includes appropriate spatial labels and landmarks to assist user orientation. • It is clear to the user how to exit the virtual environment	Darken and Sibert (1996a, 1996b), Marsh, Wright and Smith (2001), Stanney, Mollaghasemia, Reevesa, Breaux and Graeber (2003), Sutcliffe and Gault (2004), Bennett *et al.* (2010), Schofield and Lester (2010), Munoz & Chalegre (2012). Rebelo & Noriega (2012).
5	**Navigation:** • It is easy for users to move and reposition themselves in the virtual environment. • Ways of navigation are consistent throughout the system. • Logical barriers are used in areas where physical barriers are absent, but to which users should not be granted access. • Users can relocate using a terrain map.	Squires and Preece (1999), Bowman, Kruijff, LaViola and Poupyrev (2001), Stanney *et al.* (2003), Kalawsky (1999), Alessi and Trollip (2001), Munoz & Chalegre (2012), Su, Wang, Wu & Kuo (2013).
6	**Object interaction – selection and manipulation**: • Input devices are easy to use and easy to control. • Object interactions are designed realistically to reproduce real-world interaction. • The system provides the ability to rotate 3D objects and increase levels of detail when necessary for task performance.	Witmer and Singer (1998), Kalawsky (1999), Bowman *et al.* (2001), Stanney *et al.* (2003), Munoz & Chalegre (2012), Rebelo & Noriega (2012).

7	**Fidelity:** • The simulations in the system are accurate • The objects in the virtual environment move in a natural manner. • The virtual environment displays adequate levels of realism. • High-fidelity graphics are used where required.	Kalawsky (1999), Sutcliffe and Gault (2004), Bennett *et al.* (2010), Collins (2012), Schofield (2014).
8	**Various user modes:** • The system provides various user-guidance modes, e.g. Free mode, Presentation mode, Guided mode and Discovery mode. • The system provides shortcuts to frequent users.	Arendarski, Termath and Mecking (2008), Bennett *et al.* (2010).
Category 4: Context-specific criteria		
	Heuristic	**References**
1	**Authentic tasks:** • The training system supports particular work practices in the context of their natural environment. • The system is customised according to learner-specific needs and the relevance of the curriculum. • The program includes tasks applicable to the actual job context of the trainee.	Sachs (1995), Beyer and Holtzblatt (1998), Harris and Henderson (1999), Jonassen (1999), Squires and Preece (1999), Notess (2001), Reeves *et al.* (2002), Edmundson (2003), Ardito *et al.* (2004b), Chen, Toh and Fauzy (2004), Vrasidas (2004), Ssemugabi and De Villiers (2007), Teräs and Herrington (2014).
2	**Appropriate reference materials:** • The system includes additional reference materials, providing information to trainees on standard operating procedures used in the application domain. • The reference materials included in the system are relevant to the problem scenarios. • The reference materials are at a level appropriate to the trainees.	Albion (1999), Alessi and Trollip (2001), Mason, Cooper and Wilks (2015).
3	**Comprehensive scope of the system:** • The learning material in the program covers all the vital aspects relating to the topics being addressed. • The training also covers possible consequences of trainees not applying the learning material correctly in their work place.	Experience of the present researcher

4	**Adaptive design:** • The design of the training system is adaptive to changes in site practices. • The system refers to the latest current standard operating procedures. • The system randomises assessment details such as questions and multiple choice answers when presenting assessment opportunities to trainees.	Experience of the present researcher, Rusu *et al.* (2011), Su, Wang, Wu & Kuo (2013).
5	**Relevant subject matter:** • The subject matter matches the goals and objectives of the training program. • The subject matter is presented in an appropriate content structure. • The information provided in the program is accurate. • The system 'speaks the trainee's language' by using terms, phrases, symbols and concepts familiar to the trainee and common to the application domain. • The level of language use, in terms of grammar and style, is applicable to the target audience.	Nielsen (1994), Squires and Preece (1999), Alessi and Trollip (2001).
6	**Trainee preparedness:** • Trainees are shown how to use the software prior to doing the training program. • PC literacy pre-training is available to trainees not comfortable with using computers for training.	Hollender *et al.* (2010).
7	**Appropriate record keeping:** • The system maintains student records and assessment results. • The system monitors and displays student progress. • The system ensures legal compliance in the application domain by capturing detailed individual performance data.	Vrasidas (2004).
8	**Understandable and meaningful symbolic representation:** • Symbols, icons and terminology used to represent concepts and objects are used consistently throughout the program. • Symbols, icons and terminology used are intuitive within the context of the task. • Metaphors used correspond to real-world objects or concepts.	Nielsen (1994), Squires and Preece (1999), Alessi and Trollip (2001), Stanney *et al.* (2003), Dix *et al.* (2004), Oviatt (2006).

9.5. Conclusion

This chapter covered the meta-evaluation of the *DEVREF* Evaluation Framework by defining meta-evaluation, explaining the design of the meta-evaluation instrument and discussing the findings of the meta-evaluation relating to the framework's criteria and method. These findings led to several improvements to the *DEVREF* Framework and also the heuristic evaluation instrument based on the framework.

Of greater note than the improvements, however, is the fact that both the application of *DEVREF* in heuristic evaluations and the meta-evaluation indicated that the *DEVREF* Framework, as proposed in this research, is valid, comprehensive and appropriate for evaluating desktop VR training systems, and is therefore a meaningful contribution to the body of knowledge.

The formal deliverables of this chapter are the improved *DEVREF* Framework (Section 9.4) and the revised heuristic evaluation instrument, which is included as Appendix B-4. This chapter also further addressed Research Subquestion 5: "What structure, categories and criteria should be incorporated in an evaluation framework for virtual reality training systems in the mining industry" in Section 9.4. Primarily, the chapter discussed Research Subquestion 6: "How appropriate and effective is the proposed framework?" in Sections 9.2 and 9.3.

The next chapter concludes the study and revisits the research questions, makes recommendations, and mentions possible future work related to this study.

Chapter Ten

Conclusion and Recommendations

10.1. Introduction and background

The primary aim of this research, as expressed in the Main Research Question was to generate an evaluation framework to evaluate interactive desktop VR training systems for the mining industry. The rationale behind this need is explained in the next paragraph.

As stated in Section 1.3, the initial problem that led to this research was inadequate safety performance by the mining industry. Investigations pointed towards the need for possible improvement of safety training to address the problem. VR was identified as an alternative method to deliver safety training, and the potential benefits of using VR in a training environment were explained in Section 2.4. From the earliest stages of designing the VR training systems it became apparent that development of products required concomitant and rigorous evaluation. In order to assess the effectiveness of the design of such training systems, an evaluation mechanism was required that could be used to investigate the two VR prototypes designed and developed in this research. No evaluation systems were identified that targeted the unique needs of evaluating these innovative systems in the South African mining industry (Section 1.5.2), hence the need arose to design and develop an appropriate evaluation framework.

The secondary aim of the research was to propose, design and develop novel e-training interventions on the topic of safety training for mine workers operating in the underground mining environment. The training would be delivered by desktop VR technology.

To achieve the stated aims, a main research question and six subquestions were formulated. Section 10.2 revisits these research questions and summarises the findings in response to each question, while Section 10.3 details the practical, theoretical and methodological contributions of the study.

This study involved the application of design-based research based on the implementation of a new DBR process model, synthesised by the researcher. Section 10.4 reflects on how DBR was implemented and elaborates on the way in which DBR features were applied in the study. Section 10.5 discusses how the study implemented validity, reliability and triangulation. The limitations of the study are detailed in Section 10.6, while Section 10.7 gives recommendations and explores future work related to the study. Section 10.8 concludes the chapter, and the thesis.

Figure 10.1 graphically indicates the layout of this chapter.

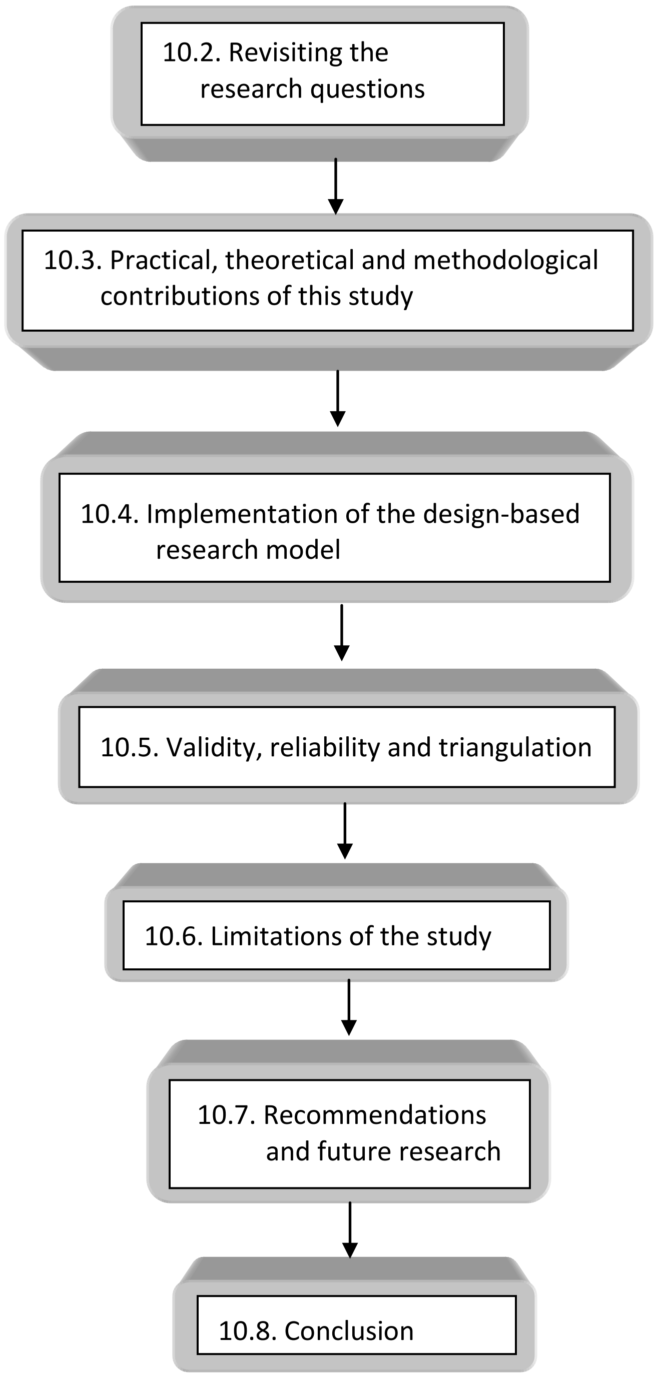


Figure 10.1: Layout of Chapter Ten.

10.2. Revisiting the research questions

As stated in Section 1.4 in Chapter One, the Main Research Question addressed by this study, is:

> **What is an appropriate and effective framework for evaluating virtual reality training systems in the mining industry?**

To answer the Main Research Question, it was subdivided into a set of multifaceted constituent subquestions, which address aspects such as the nature of VR technology, an appropriate research design for this study, and the structure for the proposed evaluation framework in terms of its categories and criteria. Furthermore, other research subquestions addressed the actual e-training systems that would be evaluated by such a framework and the development lifecycle of the systems. The final research subquestion, which was answered after the generation of the desktop VR evaluation framework (*DEVREF)*, related to the appropriateness and effectiveness of the framework itself, in other words, it is a meta-evaluative question. Six research subquestions were therefore formulated and investigated by the study. They are shown in Table 10.1, a repeat of Table 5.1. The table also indicates the section(s) in the study where each subquestion was addressed.

Table 10.1: Research subquestions with corresponding locations in thesis.

#	Research subquestion	Location in thesis	
RQ1	What is the suitability and potential of virtual reality technology for training applications in the domain of mine safety training?	Chapter 2	Sections 2.4 and 2.5
		Chapter 4	Section 4.4
RQ2	Which research paradigm is appropriate for the intended research?	Chapter 5	Sections 5.5 and 5.6
RQ3	What are the contextual requirements for virtual reality training systems for the mining industry?	Chapter 6	Section 6.5
		Chapter 7	Section 7.2
RQ4	What is an appropriate design lifecycle model for interactive desktop virtual reality training systems?	Chapter 7	Section 7.3
RQ5	What structure, categories and criteria should be incorporated in an evaluation framework for virtual reality training systems in the mining industry?	Chapter 3	Section 3.6.3
		Chapter 5	Section 5.8
		Chapter 9	Section 9.4
RQ6	How appropriate and effective is the proposed framework?	Chapter 9	Sections 9.2 and 9.3

The remainder of this section will revisit and summarise the findings of each of the six research subquestions.

10.2.1. Research Subquestion 1.

> What is the suitability and potential of virtual reality technology for training applications in the domain of mine safety training?

Section 4.4 presented statistics regarding fatality rates and serious injuries in the mining industry. They emphasise that improvements in safety training are required to reduce the human error factor. The general low literacy levels of employees in the mining sector and ineffective communication due to the range of languages spoken by the workforce, were identified as challenges to effective training.

The literature study in Chapter Two indicated that virtual reality has become a focus of research interest with regard to its potential application in training. Moreover, virtual reality-based training has been accepted in various industries as an effective form of training. VR training can be applied to a variety of workplace activities, including those of a safety-related nature. The use of VR training tools can help to reduce accidents or incidents that cause injuries and fatalities, since they present visual representations that simulate the real world and allow employees to practice skills from the safety of a computer-based simulated environment.

VR is ideal for the training of operators who perform tasks in dangerous or hazardous environments. Trainees can first practise procedures in a risk-free virtual environment, and can be exposed to 'life-threatening' scenarios in a safe and controlled situation. Section 2.4 also indicated several contexts where VR training can be used productively.

The use of virtual reality offers opportunities to the mining industry to develop tools and systems that can improve knowledge and understanding of the work environment. Virtual reality systems for the mining industry have been developed by many organisations with varying degrees of success within the areas of mine planning and design, mining equipment and training applications. This research is situated in the last-mentioned of these domains.

Virtual reality offers many possibilities in training and holds potential to increase productivity and improve safety awareness. Although the mining industry has been slow to invest in and use this advanced technology, the number of VR applications in the industry is increasing. Based on the amount of research currently undertaken on this topic, it is clear that the use of VR simulations in the mining industry will become more prevalent in the future. The hardware required to run non-immersive virtual reality systems is now available at affordable prices.

10.2.2. Research Subquestion 2

Which research paradigm is appropriate for the intended research?

Design research, in particular design-based research, was selected as the underlying research paradigm of this study because of its iterative and cyclic nature of design, evaluation and redesign, and its mandatory production of both theory and actual solutions in real-life contexts. As indicated in Section 5.5, action research could have been an alternative paradigm, since it is also iterative and can apply to inventions, interventions and products. However, DBR was deemed more appropriate for this research due to its focus on solving complex problems in context, the production of authentic artefacts, and the generation of dual outcomes, with an associated double benefit.

Section 5.6 presented the research processes used in this study and explained how a customised process model was constructed by considering preceding research models of the design and development research genre. This culminated in a newly synthesised DBR process model defined by the researcher as a foundation for this research, which was presented in Figure 5.4.

This view of DBR includes evolution of the innovation or product that is the designed solution, but also refinement of the problem, the methods and frameworks, the tools used in design and evaluation, and the design principles. The process culminates in dual outputs: (i) an implemented solution that addresses the original problem in its real-world setting and (ii) documented theory that can guide similar research and development efforts.

In Section 5.6.5 the new DBR model was compared with the DBR model of Amiel and Reeves (2008). Notable difference were:

- The new model makes provision for the design of solutions that need not be based on existing design principles, theoretical frameworks or existing knowledge bases in the problem area, as is the convention in classic DBR. In the domain of virtual reality systems, where this work resides, design theory is relatively new and established design principles are not available. Due to the innovative nature of new technologies, the new model therefore allows for conceptualisation of solutions beyond existing frameworks and principles.
- A feature of DSR (the design science variant of design research) is incorporated in the new DBR model, namely, a formal or informal proposal is advocated as an output of the first phase. The formal proposal includes a tentative design and performance criteria to evaluate the prototype. The research proposal in this DBR study preceded the research described in this thesis and is reformulated as Chapter One. Because criteria were not available at the commencement of the research, performance criteria were not part of the proposal, but were subsequently developed through an extensive literature study and presented as an evaluation framework in Section 5.8.
- The new model adapted the DBR model to include a theoretical outcome that is not merely a set of design principles. Being a design-based research methodology, the importance of design principles as an output is indeed acknowledged, but provision is also made for new theoretical contributions that extend further. Such contributions, importantly, should inform future design and evaluation in similar environments in practice. The criteria in the revised *DEVREF* Framework, presented in Section 9.4, serve both as evaluation criteria and to inform design of future desktop VR training systems.

As described in Section 5.6, four cycles of this new DBR model were applied, leading to the design, development and implementation of two VR training systems, as well as the evaluation framework to evaluate interactive desktop VR training systems for the mining industry.

The DBR process model and its application in this study was the topic of a conference paper by Van Wyk and De Villiers (2014), presented by the researcher at the South African Institute of Computer Scientists and Information Technologists' annual conference in 2014 and is included as Appendix D-3.

10.2.3. Research Subquestion 3

What are the contextual requirements for virtual reality training systems for the mining industry?

Chapter Six discussed contextual analysis for the development of virtual reality applications, applied to safety training in mines. The major context-of-use issues identified were users, tasks, equipment, the workplace environment and the training environment. The findings of the contextual analysis were discussed in Section 6.5, and the contextual analysis was also the topic of a conference paper presented at the South African Institute for Computer Scientists and Information Technologists' annual research conference in 2008, included as Appendix D-1 (Van Wyk & De Villiers, 2008).

The findings of the contextual analysis informed the design and development of the first prototype, *LSF*, which focused on generic hazards recognition and rectification. A case study was undertaken to identify specific contextual requirements for the development of the next prototype, *ISGC*, which focused on underground geological conditions and falls of ground (FOGs). The findings of the case study were described in detail in Section 7.2.

The case study findings provided valuable information for the design and development of training systems on FOGs. Not only did the case study uncover which geological conditions where present in the analysed FOGs, but also where these FOGs were located, the geological hazards present at the FOG areas, the rock dimensions of the rocks that fell, and the importance of a GCD plan to indicate the support strategy for each GCD area. This information was used to design focused content pertinent to the *ISGC* training system, as well as the computer-generated imagery to portray the simulated hazards and rock fall animations.

10.2.4. Research Subquestion 4

What is an appropriate design lifecycle model for interactive desktop virtual reality training systems?

As mentioned in Section 7.1, the researcher was unable to find an established design model directly relevant to developing and evaluating VR e-training systems for the context of this study. As this study investigates the introduction of new technology and innovative training methods into the mining industry, it was important to include all stakeholders in the development of the proposed approaches in order to gain acceptance in the industry. This emphasises the role of a design approach that takes cognisance of multiple viewpoints and information from a variety of sources.

For the design of the VR training prototypes used in this study, the interaction design lifecycle model of Rogers *et al*. (2011) was extended by the researcher to make provision for three simultaneous processes, and the subsequent integration thereof into a single product. This design model, called the extended interaction design lifecycle model, is depicted in Figure 7.15.

Application of the extended interaction design lifecycle model leads to a system evolving through a process of iterative refinement, rather than simply being developed through a linear process. Two key concepts are formative evaluation and design iteration. Formative evaluation refers to the evaluation of design ideas and aims to determine more about factors that impact on design. Design iteration allows for the refinement and revisiting of any activity within the design.

Section 7.4 describes how this extended model was applied to the design of *ISGC*. Details were given on each phase of the design and development process. Formative and summative evaluations led to several design and development improvements, and, subsequently, to an improved training system.

10.2.5. Research Subquestion 5

> What structure, categories and criteria should be incorporated in an evaluation framework for virtual reality training systems in the mining industry?

As indicated in Section 3.6.3, various factors impact on determining suitable theoretical foundations for e-learning and e-training applications. No single paradigm is appropriate for all situations, since domain, context and content all have to be considered. Technological issues and underlying educational theories should be taken into account, as well as usability, in order to find an appropriate solution that provides synergy between the learning process and interaction with the application. Specific custom-designed guidelines should be provided for the evaluation process, rather than using a set of general criteria. In line with this call, and following extensive literature reviews, this study synthesises a new set of guidelines specifically generated for evaluating VR training applications within the specific context of mining safety. The synthesised evaluation framework, called *DEVREF*, was designed for heuristic evaluation, and is presented in Section 5.8.

DEVREF consists of four categories of heuristics.

- Category 1: Instructional design – heuristics related to pedagogical effectiveness, learning theories and multimedia learning design.
- Category 2: General usability – interface design and interaction, and heuristics that support the goals of usability.
- Category 3: Virtual reality system design – heuristics specific to the design of virtual reality systems.
- Category 4: Context-related heuristics – heuristics related to the content and the application domain.

DEVREF is derived from the literature, as described in Chapter Three, as well as from the personal experience of the researcher, who has been involved for the past ten years in the design and development of virtual reality training systems for the mining industry. The iterative application of the DBR process model in this research led to the improvement of *DEVREF* and a revised version is presented in Table 9.6 in Chapter Nine. The revised framework also cites more recent sources, which confirm and extend the original framework.

10.2.6. Research Subquestion 6

How appropriate and effective is the proposed framework?

A meta-evaluation was undertaken to evaluate *DEVREF*. No standard meta-evaluation checklist was appropriate for evaluating *DEVREF,* due to its innovative and extensive nature. A custom-built meta-evaluation instrument was therefore developed by the researcher in order to evaluate both the criteria of the evaluation framework and the methodology applied for the evaluation.

DEVREF was thus scrutinised by a meta-evaluation of its criteria and methods, with the aim of assessing its quality and suggesting possible improvements. The meta-evaluation was undertaken as a systematic review by separately evaluating the methodology employed and the criteria that comprise the framework. The findings indicated that the meta-evaluation exercise was meaningful and that it achieved its aim in determining whether the evaluation statements in *DEVREF* cover all the relevant considerations. Using the feedback received, the researcher was able to improve *DEVREF*.

Chapter Eight explains how *DEVREF* was applied to evaluate the *LSF* and *ISGC* prototypes. The findings of the heuristic evaluations were presented and discussed. The interpretation of these findings, as well as the results of the meta-evaluation described in Chapter Nine, all revealed necessary improvements to the *DEVREF* Evaluation Framework. Figure 10.2 indicates the actions that led to the final version of *DEVREF.*

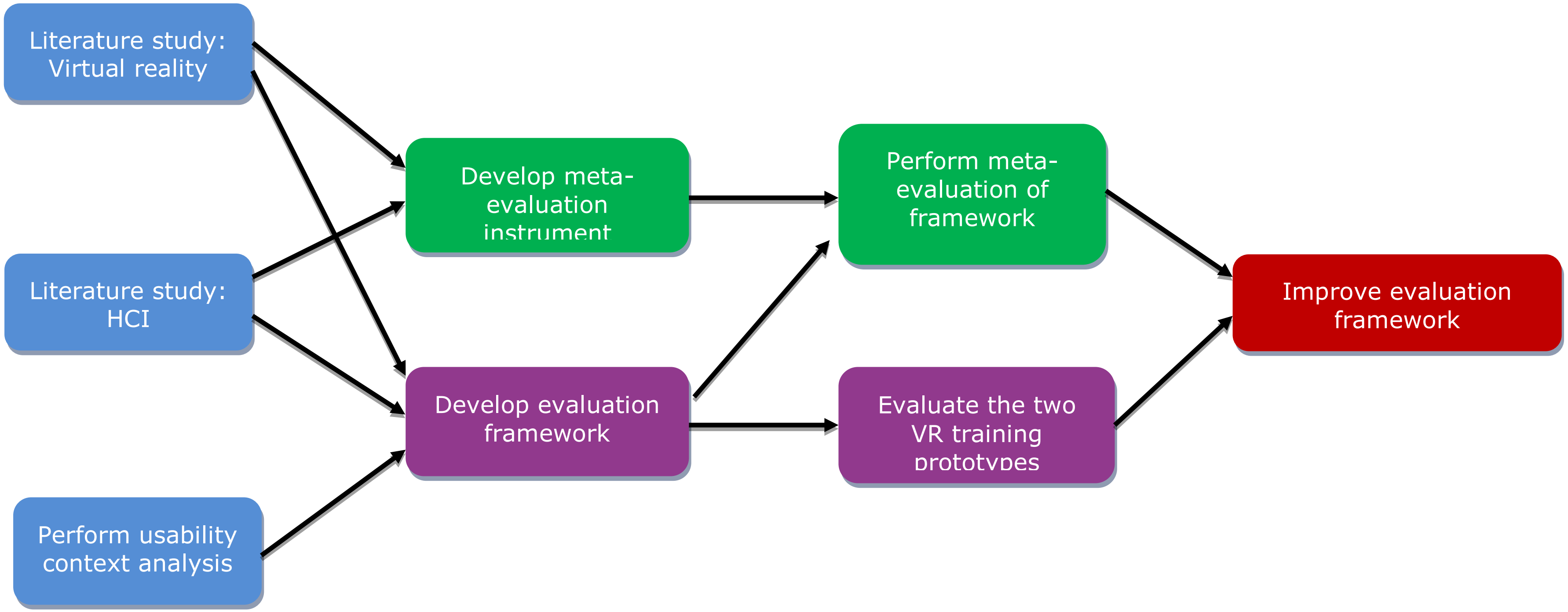


Figure 10.2: Actions leading to the final, improved evaluation framework.

10.2.7. Main Research Question

What is an appropriate and effective framework for evaluating virtual reality training systems in the mining industry?

Through the literature studies and execution of the four DBR cycles, this research led to the development and evolution of a framework that facilitates the evaluation of interactive desktop VR training systems, called *DEVREF*. Figure 1.2 presented a graphical layout of the actions taken in the research strategy of this study. Each action played a pivotal role in the evolution of *DEVREF.* Furthermore, they contributed to answering the research subquestions and thus culminated in answering the main research question. Specifically, RQ5 and RQ6 deal directly with the structure, appropriateness and effectiveness of the evaluation framework. Therefore, the final version of *DEVREF*, presented in Section 9.4, answers this main question.

10.3. Practical, theoretical and methodological contributions of this study

As described in Section 5.5.4, a DBR study has dual outcomes. These outcomes are an implemented, practical solution that addresses the original problem in its real-world setting, and a theoretical contribution in the form of design principles and/or other theory that can guide similar research and development efforts. The next two subsections highlight the practical and theoretical contributions of the study.

10.3.1. Practical contribution

An outcome of this research is a practical real-world contribution in the form of innovative desktop VR training systems. This research led to the design, development and implementation of two prototype training systems, *LSF* and *ISGC*, both currently in use as fully-fledged operational systems at Impala Platinum as part of the mine's formal training program.

The development of these authentic training systems described in this study was done in the *Centre for Creative Technologies* at the Tshwane University of Technology (TUT), under direction of the researcher. According to an agreement between TUT and a private company supplying mine safety training, STS, the intellectual property of these systems belonged to STS, and TUT would receive royalties for successful commercialisation of these systems. This agreement has led to the customised development of similar training systems for various other mines and, having moved beyond prototype stage, such interactive desktop VR systems are being used on a daily basis at fifteen mine sites throughout South Africa.

10.3.2. Theoretical contribution

As indicated in Section 5.5.4, the theoretical outcome of a DBR study can be described as contextually-sensitive design principles and theories, or as a set of design principles or guidelines that can be implemented by other researchers working in similar contexts, with the ultimate objective being the development of theory.

Being a design-based research methodology, the importance of design principles as an output is indeed acknowledged, but provision is also made for new theoretical contributions that extend further. Such contributions, importantly, should inform future design and evaluation in similar environments in practice. In this study, the synthesised DBR model, presented by the researcher, adapted the DBR process to include a theoretical outcome that is not merely a set of design principles.

This study presented an extensive list of evaluation categories and criteria, generated mainly from the literature, but also influenced by the experience of the researcher. These were integrated and structured to comprise the *DEVREF* Framework. This major theoretical contribution of the study provides a comprehensive evaluation framework for evaluating desktop VR training systems. The content of *DEVREF* could implicitly also serve as a set of design principles to guide the design of VR e-training systems.

Due to the novelty of the technology for the local mining industry and the availability and growing acceptance of desktop computer training, the *DEVREF* evaluation framework was developed specifically for desktop VR training systems. With a large, immersive training facility currently being planned at the Department of Mining Engineering at the University of Pretoria, immersive training for the local mine industry will soon be a viable prospect and the expansion of *DEVREF* to include immersive VR systems is envisaged as future work related to this study. However, due to the feasibility and recent acceptance of this non-immersive technology on desktops and the high cost of providing immersive facilities for group training, as well as the development costs of the required software, it is envisaged that desktop VR training systems will still be used at many mine training sites in the foreseeable future. This ensures the viability and future use of *DEVREF* in its current format.

10.3.3. Methodological contribution

A third contribution of this study is methodological, in that this work proposes a new DBR process model (discussed in Section 5.6.5) and an interaction design lifecycle model suitable for the design and development of VR training systems. This lifecycle model is presented in Section 7.3 and was applied in the successful design and development of the *ISGC* prototype.

10.4. Implementation of the design-based research model

This section reflects on the research design of the study. The research methodology chapter, Chapter Five, described the emergence of DBR from design science, design research and development research. DBR's characteristics were overviewed and a consolidated summary of DBR features was provided. The researcher presented a synthesised, cyclic DBR model and demonstrated how it could be applied within the process flow of a research study involving the iterative design, development, evaluation and refinement of prototype virtual reality systems for e-training in the mining industry.

The four DBR cycles presented in Section 5.6.7 detailed the application of the synthesised DBR model. To indicate how this research conforms to the DBR features presented in Table 5.3, these features are now revisited with an explanation of how each was applied to solve the problem addressed by this research. The features and explanations are provided in a tabular format in Table 10.2.

Table 10.2: DBR features applied in this study.

Feature	Application
Appropriate for complex environments	Mine safety is an important and complex issue with a sizeable effect. Even though mines have zero harm policies, more than a hundred miners die annually in work-related incidents in South Africa and thousands are injured. With more than 200 000 miners working underground in South Africa, the improvement of underground safety is an important national issue that can lead to significant results.
Problem-solving paradigm	Inadequate or insufficient training is often cited as a root cause for many mining fatalities and serious incidents. However, training outside of the direct working environment provides only limited real-life opportunities. As a result, such training may fail to make a significant impact in the tense underground working environment itself. Virtual reality-based training tools, however, can provide a basis for workers to simulate presence in

	their working conditions without the associated risks. This research explored issues related to solving a real-world practical problem and provided authentic solutions in the form of training systems, and a means of evaluating them and future artefacts.
Grounded in theory	Theory is both the foundation and an outcome of the DBR process. A theoretical framework was developed that informs future designs. Both theoretical contributions of the study, the *DEVREF* evaluation framework and the design lifecycle model, were built on, and expanded on, existing theory.
Collaborative and participative	In defining the problem and determining the context of use, interviews were held with three mine managers, two safety officers and three mine training managers. During the planning and design of the systems, the researcher collaborated with several mine training practitioners in the various cycles of design and evaluation of the systems. Furthermore, six industry and academic experts were involved in heuristic evaluation of the systems and in the meta-evaluation of the *DEVREF* evaluation framework. All these stakeholders made helpful recommendations.
Flexible and adaptable	Due to the reflection step in the DBR model applied, *DEVREF* and the VR training interventions, *LSF* and *ISGC*, were continuously developed and refined. The interventions were modified as the research progressed.
Context-sensitive	The VR training interventions functioned in an authentic setting at a large platinum mine. The *DEVREF* framework includes a category on context-specific aspects related to the mining industry, which could be adapted for evaluation of VR systems in other contexts.
Integrative	The application of *DEVREF* to the evaluation of the training interventions was the integration of theory, designed artefact, and practice in a distinct and unusual learning environment.
Innovative	This study investigated innovative ways, previously unused in the South African mining sector, of applying virtual reality technology to improve safety training. Such interactive systems have the advantage of exposing employees to numerous hazards in a safe environment, and simulate the possible consequences of unsafe acts in a 'forgiving' environment.
Iterative	Four cycles of a systematic and iterative process of analysis, design, development, evaluation and reflection were followed. The findings of each cycle continuously refined the design input to the next cycle to ensure greater relevance and enhanced performance.

Dual outcomes	This study has dual outcomes in that it contributes theory in the form of an evaluation framework with implicit design principles, as well as a practical contribution in the form of innovative, interactive VR training systems.
Pragmatic yet theoretical	Implementing the VR training systems was a substantial change in mine training practice, but the theoretical contribution is also shareable and transferable. The evaluation framework comprises four categories of criteria that can be applied in other contexts of VR training and that do not apply only to the mining industry.
Artefacts	Two authentic, tangible products were developed by following the DBR process as explained. Both these systems, *LSF* and *ISGC,* are currently in use at a number of mine training centres.
Evaluation	Each cycle of the DBR process involved rigorous and reflective inquiry to evaluate and refine the artefacts, while the fourth cycle evaluated the evaluation framework itself, which led to an improved framework.
Mixed-methods	This study employed quantitative and qualitative research methods, as indicated in the research methods section, Section 5.7.

Figure 5.4 presented the new DBR model synthesised by the researcher. This variant of DBR differs from the classic DBR cycle of Amiel and Reeves (2008) in that it extends the classic DBR model by including the design of solutions which are not necessarily based on existing design principles and technological innovations, nor are they drawn from the existing knowledge base for the problem area. Due to the innovative nature of the technology being applied, design theory is relatively new, and established evaluation frameworks are not available, making this study a pioneering effort in the South African mining industry.

10.5. Validity, reliability and triangulation

Section 5.9 discussed the generic concepts of validity, reliability and triangulation in detail. Table 10.3 presents only the aspects relevant to this study and outlines how they were implemented.

__Table 10.3: Validity, reliability and triangulation applied in the study.__

Type	Method	Application in study
Validity and Reliability of Quantitative data	Appropriate statistical instruments	Data encoding was performed using Epi-Info V6 and Stat/Transfer. Stata and SAS were used to analyse the collected data.
		Pearson's chi-square tests and Cochran-Armitage trend tests (where applicable), were conducted to compare the response patterns obtained in the *LSF* and *ISGC* evaluations. For determining statistical significance, Fisher's calculation of exact probabilities, based on the hyper-geometric distribution, were compared to the chi-square statistics.
		Internal consistency reliability was investigated by means of scale reliability tests using Cronbach alpha.
		The general linear model (GLM) approach was used to statistically verify analysis of variance.
Reliability of Qualitative data	Accuracy of interview data	Interview transcripts were transcribed verbatim without losing richness and accuracy.
		Consistent wording and context were applied in semi-structured interviews.
		One interviewer (the researcher) conducted all the interviews during the usability context analysis, thus ensuring consistency of interpretation and reducing the impact of researcher bias.

Reliability of Qualitative data (continued)	Stability of observations	The same observer (the researcher) observed the underground mining activities at various mines.
	Questionnaire volunteer bias eliminated	All participants responded under uniform conditions, namely directly after completing the prototype training.
Validity of Qualitative data	Internal validity	Surveys were pre-tested by knowledgeable colleagues to ensure that the questions measure what they are supposed to measure and that they are clear and unambiguous.
		An appropriate research methodology (DBR) was applied for answering the research questions, in an appropriate time-scale, devising and using appropriate instrumentation for data collection. Sample sizes were appropriate.
	Content validity	Responses made by the participants were used as verbatim quotations to support interpretation.
		Opinions of the researcher were indicated as such.
		The set of categories and criteria in the evaluation framework covers the domain comprehensively.
	External validity	The large size of the end-user data samples (491 participants) is representative of and applicable to the mining industry.
		Informal observations were carried out at several mines, not only at the one used for collection of the survey data.

Triangulation	Time triangulation	The same data collection instruments were used on different occasions: both prototypes were evaluated using the same heuristic evaluation framework and user satisfaction questionnaire.
		Design and subsequent improvements to the evaluation framework were informed by feedback from four DBR cycles.
	Methodological triangulation	More than one method of data collection was used.
	Participant triangulation	Different cohorts of participants were used for the two user satisfaction surveys.
	Evaluator triangulation	The six heuristic evaluators who participated were experts from different specialisation areas relevant to the study.

10.6. Limitations of the study

This research proposes virtual reality e-training interventions with the aim of improving the safety of mine workers. The impact of the e-training, however, cannot be directly related to an actual increase or decrease in injuries or fatalities among mine workers. In practice, a single individual can make a mistake leading to an incident which might cause many to be injured or killed, or no-one to be hurt, depending on the circumstances. Furthermore, not all incidents are due to worker error, as there may be a lack of management controls or falls of ground due to seismic events.

Due to the use of the VR training systems in this study as refresher training only, the actual learning of new content could not be measured and compared to other training methods. Instead, user satisfaction with the e-training was measured.

Another limitation of the study is that data collection for user satisfaction was done at one mine site only. However, Impala Platinum is the largest platinum mine site in the country, a large sample size was used (491 end-user participants for the first three DBR cycles, see Section 5.6.7) and data collection was done among different levels of employees. As mentioned in Section 10.3.1, after completion of the *LSF* and ISGC evaluations, similar VR training systems have since been developed and implemented at other mines, including gold, coal and chrome mines.

Further limitations are the high cost of immersive VR facilities, which restricts current application to desktop VR solutions, and the computer skills of the mining workforce. It is anticipated that continued exposure to this technology will raise the level of computer skills of the workforce as they interact with the VR simulations.

As indicated in Section 1.9.2, this research focuses on desktop VR and does not include the use of immersive technologies, such as cybergloves and head-mounted displays. Therefore, the evaluation framework does not include criteria related to immersive systems, but it can be expanded to include such aspects in future. Issues identified to be investigated, would be: extraneous cognitive load generated by the cognitive effort required to successfully interact with the simulation environment; usability aspects related to navigation methods and control of the equipment; as well as human factor considerations, in particular those design elements that may contribute to cybersickness.

The scope of this research relates to the use of VR in recognising hazards within the general work areas. It excludes malfunctions in safety-critical computer systems at mines. Neither does the research extend to the operation of equipment or systems, that is, the domain of simulators, but focuses on VR simulation of generic and geological hazards encountered in the underground working areas.

10.7. Recommendations and future research

The following subsections present recommendations relating to the findings of this study and explores future research options.

10.7.1. Recommendations

The application of the *DEVREF* Framework for the evaluations of the *LSF* and *ISGC* prototype systems, as well as the meta-evaluation of the framework, led to a range of improvements to *DEVREF*, which resulted in a refined framework (as described in Section 9.4). After the application and refinement of *DEVREF*, the researcher found a further 21 recent literature sources that stress the importance of many of the *DEVREF* criteria. These additional references are shown in blue in Table 9.6. This final, refined framework has not yet been applied, but has potential for future use and transfer to related domains.

The training material used in the two prototypes covered generic and geological hazards. The same design principles can be applied to other learning content, for example, generic induction, site-specific induction, equipment operation, drilling and blasting, loading and haulage, and different mining methods, to name but a few. However, the most appropriate way to present this content may differ from the instructional design of this study and the effectiveness of various instructional designs should be investigated.

To accommodate high volumes of trainees, the use of non-immersive systems on ordinary personal computers can provide a means of achieving current training goals in the South African mining industry. It is, however, expected that as the technology gradually matures in this industry, facilities will be provided for the development and implementation of more individualised systems, especially for training workers in the use of high-cost equipment, such as the continuous miner or drill rigs.

Because work underground is performed in teams, it also becomes imperative to provide a simulated training environment in which a team can work together as a unit. Semi-immersive systems using stereoscopic projection or immersive systems could cater for this need, or stereoscopic panoramic environments could be used to immerse an audience in 3D imagery. Such systems could include vision-based motion tracking systems capable of tracking and responding to movements of users. Stereoscopic projection onto a dome structure could also be used for spherical representations to cover the peripheral vision of a user standing directly in front of it. Such features would result in a truly immersive experience. Stereoscopic glasses would allow the simulations to be seen in 3D, and panoramic screen projection would enable a group of miners to experience simulations with a very high degree of realism. Trainees could be confronted with high-fidelity representations of real-world problems.

10.7.2. Future research

Although *DEVREF* was developed as an evaluation framework for the mining industry, only one category relates to the context of mining (Category 4: Context-specific criteria), where the other three categories (instructional design, generic usability and VR systems design) apply to all desktop VR training systems. This implies that *DEVREF* can easily be adapted for evaluation of VR training systems in other industries by specifying context-specific criteria related to such industries for Category 4.

Similarly, *DEVREF* can be expanded to also cover semi-immersive or immersive VR training systems by adapting criteria or adding additional criteria to Category 3 (VR systems design). With the establishment of an immersive training facility at the University of Pretoria imminent (as mentioned in Section 10.3.2), as well as the release of affordable VR headsets in the offing (as discussed in Section 2.2.1), the researcher envisages future research on extension of *DEVREF* for immersive applications.

Research topics related to group training using virtual reality could include investigating the use of audience response systems (clickers) for group interaction with virtual environments; and connecting 3D headsets in a networked environment, with each trainee wearing his own headset and seeing visuals according to his position in the virtual environment, while participating in a group exercise.

As mentioned in the Limitations section, Section 10.6, this study did not have a learning focus to demonstrate learning using VR technology, but an HCI focus on instructional design, usability, context-specific aspects, VR training systems design and evaluation. Future research could investigate actual learning of content, compared to, or in conjunction with, other training methods. Specifically, learning using immersive virtual environments would be important future research related to this study.

Moreover, the continual development and implementation of VR training systems in the local mining industry is gradually raising the maturity level of the use of visualisation and simulation in training. This will allow for more complex applications of the technology in future, for example, a current international trend is the use of building information modelling (BIM). BIM represents the development and use of VR technology to simulate the planning, design, construction and operation of a facility. It helps architects, engineers and managers to visualise what is to be built in a simulated environment and to identify potential design, construction or operational problems. Applying BIM to the mining industry implies the 3D visualisation of the mine design, construction and operation throughout its simulated lifetime, linking production and financial aspects to the dynamic visuals. The application of BIM to the mine industry is a new research area (Gomez, 2012; Howe, 2014; Sundt Connections, 2013), and will require complex VR environments.

Due to the poor literacy levels of the general workforce in the South African mining industry, a more intuitive means of interaction should be investigated to replace some of the on-screen text in the training systems.

In Section 2.3.3, augmented reality (AR) was defined as the overlaying of computer-generated imagery or data onto real-world imagery. AR technology makes it possible to project data, diagrams, animation or video onto transparent glasses, which the user can then see while viewing the real world. The prototypes in this study did not make use of AR, but it can potentially be used in future visualisation training systems, for example, projection of diagrams of machine parts or instruction manuals on goggles to assist workers while making repairs on the actual machine. Further research on this topic can lead to the expansion of *DEVREF* to include criteria relevant to AR.

10.8. Conclusion

This chapter presented a summary of the work undertaken in this study. Following a brief introduction, the Main Research Question and the six subquestions that this study aimed to answer, were revisited and answered one by one. Thereafter the practical, theoretical and methodological contributions of this study were presented.

Section 10.4 reflected on the implementation of DBR, and Section 10.5 discussed how the study implemented validity, reliability and triangulation. The limitations of the study were detailed in Section 10.6, while Section 10.7 presented recommendations and future work related to the study.

In summary, this study described the application of design-based research for the design and development of desktop virtual reality systems for safety training in the South African mining industry. The process flow of the research moved from a complex real-world problem to dual outcomes, namely a practical real-world solution in the form of two virtual reality training systems and a contribution to documented theory in the form of an evaluation framework. This desktop VR evaluation framework (*DEVREF*) can be applied to evaluate the effectiveness of the design of desktop VR systems regarding instructional design, usability, VR systems design and context-specific criteria in the mining industry. Furthermore, this study proposed a new process model for DBR and an interaction design lifecycle model for VR training systems. All these artefacts should provide great value to training in the South African mining industry.

The findings of this study suggest that the DBR approach using the proposed new DBR model, is appropriate for designing and evaluating VR training artefacts to enhance mine safety training. As a *practical* contribution, the two systems are in use at several mine training centres. As a *theoretical* contribution, the evaluation framework is transferable and customisable to other industry contexts. Furthermore, the criteria in the framework also serve as design principles. Finally, the proposed new DBR process model and the extended interaction design lifecycle model both provide *methodological* contributions.

This study made a major contribution to the body of knowledge on the design, development and evaluation of interactive desktop VR training systems for the South African mining industry.

References

5DT. 2012. *Fifth Dimension Technologies: Virtual Reality for the Real World* [Online]. Available from: http://www.5dt.com/products.html [Accessed: 29/11/2012].

ACM (Association for Computing Machinery). 1996. *ACM SIGCHI Curricula for Human-Computer Interaction* [Online]. Available from: http://old.sigchi.org/cdg/cdg2.html#2_1 [Accessed: 11/02/2012].

ActiveWorlds. 2012. *Active Worlds and Education* [Online]. Available from: http://www.activeworlds.com/edu/index.asp [Accessed: 17/03/2012].

Adams, D., du Plessis, A.G., Gumbie, A. & Willis, R.P.H. 2007. *Introduction to Safety Practice in South African Mines*. Mine Health and Safety Council, South Africa.

Adebesin, F., Kotze, P. & Gelderblom, H. 2010. The Complementary Role of Two Evaluation Methods in the Usability and Accessibility Evaluation of a Non-Standard System. *Proceedings of the Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT 2010).*

Adikari, S., McDonald, C. & Campbell, J. 2011. A Design Science Framework for Designing and Assessing User Experience. In: *J.A. Jacko (ed). Human-Computer Interaction, Part 1. HCI1*. LNCS 6761:25-34. Berlin: Springer-Verlag.

Albion, P.R. 1999. *Heuristic Evaluation of Multimedia: From Theory to Practice*. [Online]. Available from: http://www.usq.edu.au/users/albion/papers/ascilite99.html [Accessed: 27/01/2009].

Alessi, S.M. & Trollip, S.R. 2001. *Multimedia for Learning: Methods and Development*. 3rd ed. Massachusetts: Allyn & Bacon.

Alonso-Rios, D., Vazquez-Garcia, A., Mosquiera-Rey, E. & Moret-Bonillo, V. 2010. A Context-of-Use Taxonomy for Usability Studies. *International Journal of Human-Computer Interaction,* 26(10):941–970.

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

AMMSA. 2012. *Association of Mine Managers of South Africa – Safety and Health* [Online]. Available from: http://www.ammsa.org.za/aa/safety_health.asp [Accessed: 05/05/2012].

Anderson, T. & Shattuck, J. 2012. Design-based research: A decade of progress in education research? *Educational Researcher, 41*(1):16-25.

Ardito, C., De Marsico, M., Lanzilotti, R., Levialdi, S., Rossano, V. & Tersigni, M. 2004a. Usability of E-Learning Tools. *Proceedings of the working conference on advanced visual interfaces*. Gallipoli, Italy:80-84.

Ardito, C., Costabile, M.F., De Marsico, M., Lanzilotti, R., Levialdi, S., Plantamura, P., Roselli, T., Rossano, V. & Tersigni, M. 2004b. Towards Guidelines for Usability of e-Learning Applications. *User-Centered Interaction Paradigms for Universal Access in the Information Society - 8th ERCIM Workshop on User Interfaces for All 2004,* 185-202. Berlin: Springer-Verlag.

Ardito, C., Costabile, M.F., De Marsico, M., Lanzilotti, R., Levialdi, S., Roselli, T. & Rossano, V. 2006. An Approach to Usability Evaluation of e-Learning Applications. *Information Society*, 4(3):270-283.

Arendarski, B., Termath, W. & Mecking, P. 2008. Maintenance of Complex Machines in Electric Power Systems Using Virtual Reality Techniques. *Proceedings of the 2008 IEEE International Symposium on Electrical Insulation ISEI 2008.* Vancouver.

Avouris, N.M. 2001. An introduction to software usability, *Proceedings of the 8th Panhellenic Conference on Informatics*, 2:514-522.

Babbie, E. 2010. *The Basics of Social Research.* 3rd ed. Thomson Whadsworth, Canada.

Badenhorst, W. 2011. *Mining Industry Under Pressure to Reduce Fatalities* [Online]. Available from: http://www.miningsafety.co.za/newscontent/89/Mining-industry-under-pressure-to-reduce-fatalities [Accessed: 03/12/2014].

Baker, D. 2006. Interview with Safety, Health and Environment Manager, Anglo Platinum Bafokeng-Rasimone. (Notes in possession of author).

Bamberger, M.J., Rugh, J. & Mabry, L. 2006. *Real World Evaluation: Working under Budget, Time, Data, and Political Constraints*, Sage Publications.

Bannan-Ritland, B. 2003. The Role of Design in Research: The integrative learning design framework. *Educational Researcher*, 32(1):21-24.

Barab, S.A. & Duffy, T.M. 2000. From Practice Fields to Communities of Practice. In: *D.H. Jonassen & S.M.Land (eds). Theoretical Foundations of Learning Environments.* NJ:Lawrence Erlbaum.

Barab, S., Sadler, T., Heiselt, C., Hickey, D. & Zuiker, S. 2007. Relating Narrative, Inquiry, and Inscriptions: A Framework for Socioscientific Inquiry. *Journal of Science Education and Technology*, 16:59-82.

Barab, S. & Squire, K. 2004. Design-based Research: Putting a Stake in the Ground. *The Journal of the Learning Sciences*, 13(1):1-14.

Barber, C. 2002. *Heuristics for Websites* [Online]. Available from: http://www.id-book.com/catherb/Website_heurs.php [Accessed: 13/07/2008].

Barisic, A., Amaral, V., Goulao, M. & Barroca, B. 2012. Evaluating the Usability of Domain-specific Languages. In: *M. Mernik (ed). Formal and Practical Aspects of Domain-Specific Languages: Recent Developments*, 2012:386–407. IGI Global.

Barry, B. 1995. Shifting the Balance: the Leon Commission of Enquiry into Health and Safety in the Mining Industry. *SA Labour Bulletin*, 19(2):65-70.

Belkhiter, N., Boulet, M., Baffoun, S. & Dupuis, C. 2003. Usability Inspection of the ECONOF System's User Interface Visualization Component. In: *C. Ghaoui. (ed), Usability Evaluation of Online Learning Programs.* Hershey, PA: Information Science Publishing.

Bell, P. 2004. On the theoretical breadth of design-based research in education. *Educational Psychologist*, 39(4):243−252.

Bennett, L., Stothard, P. & Kehoe, J. 2010. Evaluating the Effectiveness of Virtual Reality Learning in a Mining Context. In: *D.E. Leigh (ed). SimTecT 2010 Conference Proceedings.* Simulation Industry Association of Australia. Brisbane Australia.

Berkley, J.J. 2003. *Haptic Devices* [Online]. Available from: http://www.hitl.washington.edu/people/tfurness/courses/inde543/READINGS-03/BERKLEY/White%20Paper%20-%20Haptic%20Devices.pdf [Accessed: 09/01/2014].

Bevan N. 1997. Quality and usability: A new framework, In: *E. van Veenendaal & J. McMullan (eds). Achieving software product quality*, Tutein Nolthenius, Netherlands.

Beyer, H. & Holtzblatt, K. 1998. *Contextual design: Defining customer-centered systems*. San Francisco: Morgan Kaufmann.

Bias, R.G. 1994. The Pluralistic Usability Walkthrough: Coordinated Empathies. In: *J. Nielsen & R.L. Mack (Eds). Usability Inspection Methods*. New York: John Wiley & Sons.

Biffi, M. 2000. *A baseline survey to assess health and safety research needs of smaller, previously uncontrolled mines other than gold, platinum and coal mines employing fewer than 300 people* [Online]. Available from: http://www.simrac.co.za/report/Reports/thrust9/oth605/oth605.htm [Accessed: 11/11/07].

Bise, C.J. 1997. Virtual Reality: Emerging Technology for Training of Miners. *Mining Engineering,* 49(1).

Black, E. 1995. *Behaviorism as a Learning Theory* [Online]. Available from: http://129.7.160.115/inst593/Behaviorism.htm [Accessed: 02/11/2008].

Blandford, A., Keith, S., Connell, I. & Edwards, H. 2004. Analytical Usability Evaluation for Digital Libraries: A Case Study. *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries.* 2004:27-36. Tucson: ACM Press.

Boghossian, P. 2006. Behaviorism, Constructivism and Socratic Pedagogy. *Educational Philosophy and Theory,* 38(6):713-722.

Bower, K.M. 2003. *When to use Fisher's Exact Test* [Online]. Available from: http://www.minitab.com/uploadedFiles/Content/News/Published_Articles/fisher_exact_test.pdf [Accessed: 14/05/2014].

Bowler, L. & Large, A. 2008. Design-based research for LIS. *Library & Information Science Research*, 30(2008):39–46.

Bowling, S.R., Kaewkuekool, S., Khasawneh, M.T., Jiang, X. & Gramopadhye, A.K. 2003. Evaluating the use of Virtual Reality for training in aircraft inspection. *Proceedings of the 8th Annual International Conference on Industrial Engineering – Theory, Applications and Practice*, Las Vegas, Nevada, USA, Nov:248.

Bowman, D.A., Gabbard, J.L. & Hix, D. 2002. A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods. *Presence: Teleoperators and Virtual Environments*, 11(4):404–424.

Bowman. D.A., Koller, D. & Hodges, L.F. 1998. A methodology for the evaluation of travel techniques for immersive virtual environments. *Virtual Reality,* 3(2):120-131.

Bowman, D.A., Kruijff, E., LaViola, J.J. & Poupyrev, I. 2001. An Introduction to 3-D User Interface Design. *Presence: Teleoperators and Virtual Environments*, 10(1):96–108.

Bowman, D.A. & McMahan, R.P. 2007. Virtual Reality: How Much Immersion Is Enough? *Computer*, 40(7):July 2007.

Briggs, G. 2014. Creating SA Mining Inc. *Presentation at AMMSA Annual General Meeting* [Online]. Available from: http://www.ammsa.org.za/knowledge-hub/presentation-archive/2014/finish/27-2014/1191-creating-sa-mining-inc-graham-briggs [Accessed: 05/11/2014].

Brooke, J. 1996. SUS: a "quick and dirty" usability scale. Chapter from *Usability Evaluation in Industry:* 189-194. London: Taylor & Francis.

Brooke, J. 1998. *System Usability Scale* [Online]. Available from: www.usabilitynet.org/trump/documents/Suschapt.doc [Accessed: 12/05/2008].

Brown, A.L. 1992. Design experiments: Theoretical and methodological challenges in creating complex interventions. *Journal of Learning Sciences, 2*(2):141-178.

Brown, J. 2009. *Miners are still getting shafted on mine safety* [Online]. Available from: http://www.busrep.co.za/index.php?fSectionId=3998&fArticleId=5151063 [Accessed: 14/05/2010].

Brown, J. 2010. *Mine safety law should be given sharp teeth* [Online]. Available from: http://www.busrep.co.za/index.php?fArticleId=5448621 [Accessed: 15/052010].

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

Brown, M., Sharples, S. & Harding, J. 2013. Introducing PEGI: A Usability Process for the Practical Evaluation of Geographic Information. *International Journal of Human-Computer Studies*, 71(2013):668–678.

Bruner, J. 1990. *Constructivist Theory* [Online]. Available from: http://www.gwu.edu/~tip/bruner.html [Accessed: 13/07/2009].

Brunken, R., Plass, J.L. & Leutner, D. 2004. Assessment of Cognitive Load in Multimedia Learning with Dual-Task Methodology: Auditory Load and Modality Effects. *Instructional Science,* 32:115-132.

Bryson, S. 1996. Virtual Reality in Scientific Visualization. *Communication of ACM*, 39(5):62-71.

Burdea, G., Patounakis, G. & Popescu, G. 1998. Virtual reality training for the diagnosis of prostate cancer. *Proceedings of the Virtual Reality Annual International Symposium*, Atlanta, GA, 190-197.

Bureau of Labor. 1999. *Number, Percent and Rate of Fatal Occupational Injuries* [Online]. Available from: http://www.bls.gov/iif/oshwc/cfoi/cftb0092.pdf [Accessed: 10/01/2007].

Burgess-Limerick, R. & Steiner, L. 2006. Injuries Associated with Continuous Miners, Shuttle Cars, Load-Haul-Dump, and Personnel Transport in New South Wales Underground Coal Mines. *Mining Technology,* 115:160-168.

Burkle, M. & Kinshuk, K. 2009. Learning in virtual worlds: The challenges and opportunities. *IEEE 2009 International Conference on CyberWorlds*, UK, 320-327.

Business Report. 2008. *Mine safety a collective responsibility* [Online]. Available from: http://www.busrep.co.za/index.php?fArticleId=4379207&fSectionId=552&fSetId=662 [Accessed: 15/05/2010].

Business Report. 2009. *Safety knocks Gold Fields back R2.29bn* [Online]. Available from http://www.busrep.co.za/index.php?fSectionId=563&fArticleId=4816982. [Accessed: 30/07/2009].

Canqun, H. & Zhonghua, S. 2010. The Principles of Human-Computer Interface Design in the Multimedia Courseware. *International Conference on Artificial Intelligence and Education (ICAIE)*, Hangzhou, October 2010, 69-72.

Carroll, J.M. 2000. *Making Use: Scenario based design for Human Computer Interaction*. Cambridge, MA:MIT Press.

CAT, 2014. *CAT Simulators: Operator Training, Virtual Technology, Real Results* [Online]. Available from: http://www.cat.com/en_US/support/operations/operator-training1/simulators.html [Accessed: 10/01/2015].

Cedefop. 2002. *E-learning and training in Europe*. European Centre for the Development of Vocational Training. Luxembourg: Office for Official Publications of the European Communities, 5.

Chadwick, J. 2009. Simulators and Virtual Training. *International Mining September 2009* [Online]. Available from: http://www.infomine.com/publications/docs/InternationalMining/Chadwick2009s.pdf [Accessed: 22/04/2010].

Chakraborty, P.R. & Bise, C.J. 2000. A Virtual Reality-based Model for Task-Training of Equipment Operators in the Mining Industry. *Mineral Resources Engineering*, 9(4):437-449.

Chalmers, P.A. 2003. The Role of Cognitive Theory in Human-Computer Interface. *Computers in Human Behavior*, 19:593-607.

Chareen, S. 2007. Micro-Level Design for Multimedia-Enhanced Online Courses. *Journal of Online Learning and Teaching*, 3(4):383–394.

Chen, C. 2004. *Information Visualization: Beyond the horizon*. 2nd ed. Berlin: Springer-Verlag.

Chen, C. J. 2010. *Virtual reality (VR)-based learning environment: Design, develop, evaluate*. Saarbrücken, Germany: Lambert Academic.

Chen, C.J., Toh, S.C. & Fauzy, W.M. 2004. The Theoretical Framework for Designing Desktop Virtual Reality-Based Learning Environments. *Journal of Interactive Learning Research,* 15(2):147-167. Norfolk, VA: AACE.

Citizen. 2009. *Mining Deaths up on last year: Solidarity* [Online]. Available from: http://www.citizen.co.za/index/article.aspx?pDesc=109768,1,22 [Accessed: 29/01/2010].

Clark, R.C. & Mayer, R.E. 2003. *E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning.* San Francisco, CA: Pfeiffer Publishing.

Clark, R.C. & Mayer, R.E. 2011. *Proven Guidelines for Consumers and Designers of Multimedia Learning*. San Francisco: John Wiley & Sons.

Clark, R.C. & Taylor, D. 1994. The Causes and Cures of Learner Overload. *Training*, 31(7):40-43.

Cobb, P. 2001. Supporting the improvement of learning and teaching in social and institutional context. *In: S. Carver & D. Klahr (eds). Cognition and instruction: Twenty-five years of progress:* 455–478. Cambridge, MA: Lawrence Erlbaum.

Cockton, G. 2013. Usability Evaluation. In: *M. Soegaard & R.F. Dam (eds). The Encyclopedia of Human-Computer Interaction*, 2nd ed. Aarhus, Denmark: The Interaction Design Foundation [Online]. Available from: https://www.interactiondesign.org/encyclopedia/usability_evaluation.html [Accessed: 04/12/2014].

Cohen, L., Manion, L. & Morrison, K. 2011. *Research Methods in Education*. 7th ed. London: Routledge.

Coka, B. 2012. Mining: SA's Economic Lifeline Threatens to Underperform [Online]. Available from: http://www.productivitysa.co.za/index.php/talking-productivity-sections/mining-sa-s-economic-lifeline-threatens-to-underperform [Accessed: 05/12/2014].

Collins, A. 1992. Toward a design science of education. *In: E. Scanlon & T. O'Shea (Eds.), New directions in educational technology:* 15-22. New York: Springer-Verlag.

Collins, A., Joseph, D. & Bielaczyc, K. 2004. Design research: Theoretical and methodological issues. *The Journal of the Learning Sciences, 13*(1):15-42.

Collins, C. 2012. Should virtual environments be realistic? Hypergrid Business Magazine, May 2012 [Online]. Available from: http://www.hypergridbusiness.com/2012/05/should-virtualenvironments-be-realistic/ [Accessed: 16/01/2015].

CoM (Chamber of Mines of South Africa). 2004. *Annual Report 2004* [Online]. Available from: http://www.chamberofmines.org.za/media-room/mining-publications [Accessed: 10/01/2007].

CoM. 2007. Chamber of Mines - *The South African Mining Industry's Sustainability and Transformation Report 2007* [Online]. Available from: http://www.bullion.org.za/Departments/SafetySusDevl/Downloads/SDRtext07.pdf [Accessed: 09/06/2010].

CoM. 2010b. *Chamber of Mines – Platinum group metals in South Africa* [Online]. Available from**:** http://www.bullion.org.za/Education/Platinum.htm [Accessed: 22/05/2010].

CoM. 2012a. *Chamber of Mines* [Online]. Available from: http://www.bullion.org.za/ [Accessed: 22/05/2012].

CoM. 2012b. *Chamber of Mines – Platinum group metals in South Africa* [Online]. Available from: http://www.bullion.org.za/Education/Platinum.htm [Accessed: 22/05/2012].

CoM. 2013. *Chamber of Mines Facts and Figures 2013* [Online]. Available from: http://chamberofmines.org.za/media-room/facts-and-figures [Accessed: 12/09/2014].

Costabile, M.F., De Marsico, M., Lanzilotti, R., Plantamura, V.L. & Roselli, T. 2005. On the Usability Evaluation of E-Learning Applications. In: *Proceedings of the 38th Hawaii*

International Conference on System Science, 2005:1-10. Washington: IEEE Computer Society.

Coursaris, C.K. & Kim, D.J. 2007. A research agenda for mobile usability. *Proceedings of the 2007 Conference on Human Factors in Computing Systems*, 2007:2345–2350.

Cramer, H.S.M., Evers, V., Zudilova, V. & Sloot, P.M.A. 2004. Context analysis to support development of virtual reality applications. *Virtual Reality*, 7(2004):177–186.

Cresswell, J.W. 2009. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches.* 3rd ed. California, USA: Sage Publications.

Daden, 2014. *Virtual Reality for Training, Learning, Collaboration and Visualisation* [Online]. Available from: www.daden.co.uk [Accessed: 09/01/2015].

Darken, R.P. & Sibert, J.L. 1996a. Navigating in large virtual worlds. *The International Journal of Human-Computer Interaction*, 8(1):49–72.

Darken, R.P. & Sibert, J.L. 1996b. Wayfinding strategies and behaviors in large virtual environments. *Human Factors in Computing Systems, CHI '96 Conference Proceedings:* 142–149, Vancouver.

Davids, F. 1997. A new generation of qualified miners emerges. *Mail & Guardian Newspaper*, February 21. Johannesburg.

Dawson, D.L., Meyer, J., Lee, E.S. & Pevec, W.C. 2007. Training with simulation improves residents' endovascular procedure skills. *Journal of Vascular Surgery*, 45(1):149-154.

DBRC (Design-based Research Collective). 2003. Design-based Research: An Emerging Paradigm for Educational Inquiry. *Educational Researcher*, *32*(1):5-8.

DBRC (Design-based Research Collective). 2006. *Design-based Research for Advancing Educational Technology*, EDIT 9990 [Online]. Available from: http://it.coe.uga.edu/~treeves/EDIT9990/EDIT99909Jan06.ppt [Accessed: 16/03/2011].

DBRSIG (Design-based Research Special Interest Group). 2006. *What is Design-based Research?* [Online]. Available from: http://cider.athabascau.ca/CIDERSIGs/DesignBasedSIG/ [Accessed: 29/03/2011].

Dede, C. 2005. Why Design-Based Research is Both Important and Difficult. *Educational Technology*. Jan-Feb 2005.

De Kock, E., Van Biljon, J. & Pretorius, M. 2009. Usability evaluation methods: Mind the gaps. *Proceedings of SAICSIT'09*, ACM, Vanderbijlpark, South Africa, 122-131.

Delabbio, F.C., Hitchcock, S., Iturregui, L., Dunn, P.G. & Eastick, D. 2002. The impact of 3D CAD visualization and virtual reality on interpretation and development strategies for mining and mineral processing equipment design, operation, and safety. *Proceedings of 5th Northern American Rock Mechanics Symposium (NARMS-TAC)*, Toronto.

Delabbio, F.C., Dunn, P.G., Iturregui, L. & Hitchcock, S. 2003. The application of 3D CAD visualization and virtual reality in the mining & mineral processing industry. *HATCH Mining Technology Unit* [Online]. Available from: http://www.hatch.ca/Mining_Mineral_Processing/Articles/3D_CAD_Visualization_Virtual_reality_in_MMP.pdf [Accessed: 13/04/2010].

Denby, B. & Schofield, D. 1999. Role of Virtual Reality in Safety Training of Mine Personnel. *Mining Engineering*, October:59-64.

Denby, B., Schofield, D., McClarnon, D.J., Williams, M.W. & Walsha, T. 1998. Hazard Awareness training for mining simulations using 'virtual Reality'. APCOM 98 – Applications of Computers and Mathematics in the Minerals Industries. *Proceedings of the 27th APCOM International Symposium*. London.

Dennis, K.A., & Harris, D. 1998. Computer-based simulation as an adjunct to ab initio flight training. *International Journal of Aviation Psychology,* 8(3):261-276.

Department of Minerals and Energy. 2008. *SAMRASS* [Online]. Available from: http://www.dme.gov.za/pdfs/mhs/samrass/cb_sect_e.pdf [Accessed: 16/06/2008].

De Strulle, A. 2004. Differentiation of the Causal Characteristics and Influences of Virtual Reality and the Effects on Learning at a Science Exhibit. *PhD Thesis*. University of San Diego.

De Villiers, M.R. 2005a. e-Learning artefacts: Are they based on learning theory? *Alternation,* 12(1b):345-371.

De Villiers, M.R. 2005b. Interpretive research models for Informatics: action research, grounded theory, and the family of design- and development research. *Alternation,* 12(2):10-52.

De Villiers, M.R. 2012. Models for Interpretive Information Systems Research, Part 2: Design research, development research, design-based research, design-science research – a meta-study and examples. In: *M. Mora, O. Gelman, A. Steenkamp & M.S. Raisinghani (eds). Research Methodologies, Innovations and Philosophies in Software Systems Engineering and Information Systems:* 238-255. IGI Global.

De Villiers, M.R. & Harpur, P.A. 2013. Design-based research – the educational technology variant of design research: Illustrated by the design of an m-learning environment. *Proceedings of SAICSIT '13*, October 2013, East London, South Africa.

Dijkstra, S. 2001. The Design Space for Solving Instructional-Design Problems. *Instructional Science*, 29(2):275-290.

Dinsmore, M., Langrana, G., Burdea, G. & Ladeji, J. 1997. Virtual reality training for palpation of subsurface tumors. *Proceedings of the Virtual Reality Annual International Symposium*, Albuquerque, NM. 1997:54-60.

Disney World, 2013. *DisneyQuest Indoor Interactive Theme Park* [Online]. Available from: http://disneyworld.disney.go.com/destinations/downtown-disney/entertainment/disney-quest-indoor-interactive-theme-park/ [Accessed: 11/03/2013].

Dix, A.J., Finlay, J.E., Abowd, G.D. & Beale, R. 2004. *Human Computer Interaction,* 3rd Edition. Harlow Assex: Pearson.

DME (Department of Mineral and Energy Affairs). 2005. *Guideline for Enforcement of the Mine Health and Safety Act* [Online]. Available from: http://www.dme.gov.za/publications/pdf/guidelines/guideline_enforcement_mhs_act.pdf [Accessed: 20/07/2007].

DMR. 2010a. *Department of Mineral Resources – Accident Statistics* [Online]. Available at: http://www.dme.gov.za/mhs/accident_stats.stm [Accessed 24/05/2010].

DMR. 2010b. *Department of Mineral Resources – Mine Health and Safety key documents* [Online]. Available at: http://www.dme.gov.za/mhs/documents.stm [Accessed 24/05/2010].

DMR. 2012. *Department of Mineral Resources - Occupational Safety* [Online]. Available at: http://www.dmr.gov.za/mine-health-a-safety/occupational-safety.html [Accessed 24/09/2013].

DMR. 2013. *Department of Mineral Resources Annual Report 2012/2013* [Online]. Available at: http://www.pmg.org.za/report/20131022-department-mineral-resources-201213-annual-report-briefing [Accessed 12/05/2014].

DMR. 2014a. *Department of Mineral Resources - Accident Statistics* [Online]. Available at: http://www.dmr.gov.za/mine-health-a-safety.html [Accessed 12/05/2014].

DMR. 2014b. *Department of Mineral Resources Annual Report 2013/2014* [Online]. Available at: http://www.dmr.gov.za/publications/annual-report/summary/26-annual-reports/1219-annual-report-20132014.html [Accessed 02/12/2014].

DMR. 2014c. *Department of Mineral Resources – Mine Managers Government Certificate of Competency* [Online]. Available at: http://www.dmr.gov.za/mine-managers.html [Accessed 03/05/2014].

Dringus, L.P. & Cohen, M.S. 2005. An Adaptable Usability Heuristic Checklist for Online Courses. *Proceedings of 35th ASEE/IEEE Frontiers in Education Conference*, October 2005, Indianapolis.

DSK, 2010. *Deutsche Steinkohle Products* [Online]. Available from: http://www.rag-deutsche-steinkohle.de/ [Accessed: 28/04/2010].

DST. 2010. *Department of Science and Technology - Safety in Mines Research Advisory Committee* [Online]. Available from: http://www.dst.gov.za/s-t-landscape/S-T%20Funding%20Agencies/SIMRAC. [Accessed: 12/03/2010].

Dua, T. 2014. *Volvo launches virtual reality campaign on Google Cardboard* [Online]. Available from: http://digiday.com/brands/volvo-becomes-first-automaker-use-google-cardboard-offers-test-drives-new-car/ [Accessed: 10/01/2015].

Duffy, T.M. & Jonassen, D.H. 1991. Constructivism: New Implications for Instructional Technology? *Educational Technology*, 31(5):7-12.

Dumas, J.S. 2003. User-based Evaluations, In: *J.A. Jacko & A. Sears (eds). The Human Computer Interaction Handbook.* Mahwah: Lawrence Erlbaum.

Dumas, J.S. & Fox, J. 2007. Usability testing: Current practice and future directions. In: *A. Sears & J.Jacko (eds). The Human Computer Interaction Handbook*, 2nd ed, 1129-1149. New York: Lawrence Erlbaum Associates.

Dumas, J.S. & Reddish, J.C. 1999. *A practical guide to usability testing*. Exeter:Intellect.

EADS, 2006. EADS Military Air Systems. *European Aeronautics Defence and Space Company* [Online]. Available from: http://www.eads.net/1024/es/businet/defence/mas/technologies/virtual_reality/virtual_reality.html [Accessed: 30/03/2007].

Edelson, D.C. 2002. Design Research: What we learn when we engage in design. *Journal of the Learning Sciences*, 11(1):105-121.

Edmundson, A. 2003. *Decreasing Cultural Disparity in Educational ICTs: Tools and Recommendations* [Online]. Available from: http://tojde.anadolu.edu.tr/tojde11/articles/edmundson.htm *Turkish Online Journal of Distance Education*, 4(3). [Accessed: 12/07/2009].

Emory, 2004. Researchers Say Virtual Reality Training Is Wave of the Future for Cardiovascular Medicine. *Emory University Hospital* [Online]. Available from: http://whsc.emory.edu/press_releases2.cfm?announcement_id_seq=2959 [Accessed: 23/03/2010].

Endsley, M.R. & Jones, D.G. 2012. *Designing for Situation Awareness – An Approach to User-Centred Design*. 2nd ed. Florida: CRC Press.

EPRI (Economic Policy Research Institute). 2003. Virtual reality construction for cost reduction of advanced nuclear plants. *Strategic Science and Technology,* September 2003.

Ertmer, P.A. & Newby, T.J. 2013. Behaviorism, Cognitivism, Constructivism: Comparing Critical Features from an Instructional Design Perspective. *Performance Improvement Quarterly*, 26(2), 2013:43-71, reprinted from *Performance Improvement Quarterly*, 6(4), 1993:50–72.

Estrella, P., Popescu-Belis, A. & Underwood, N. 2005. Finding the system that suits you best: towards the normalization of MT evaluation. *Proceedings of the 27th International Conference on Translating and the Computer (ASLIB)*: 23–34.

Evangelho, J. 2014. *Meet Sinister, A Haptic PC Gaming Device With Embedded Vivitouch Technology* [Online]. Available from: http://www.forbes.com/sites/jasonevangelho/2014/01/07/meet-sinister-a-haptic-pc-gaming-device-with-embedded-vivitouch-technology/ [Accessed: 09/01/2014].

Fasel, J.H., Gingins, P., Kalra, P., Magnenat-Thalmann, N., Baur, C., Cuttat, J.F., Muster, M. & Gailloud, P. 1997. Liver of the "visible man". *Clinical Anatomy,* 10:389-393.

Fernandez, N. 2010. Interview with Group Chief Rock Engineer, Impala Platinum. (Notes in possession of author).

Ferwerda, J. 2003. Three varieties of realism in computer graphics. *Proceedings of SPIE Human Vision and Electronic Imaging*, 290-297.

Filigenzi, M.T., Orr, T.J. & Ruff, T.M. 2000. Virtual Reality for Mine Safety Training, *Applied Occupational and Environmental Hygiene,* 15(6).

Foltz, C., Schneider, N., Kausch, B., Wolf, M., Schlick, C. & Luczak, H. 2008. Usability Engineering. *Collaborative and Distributed Chemical Engineering*, 527-554. Berlin: Springer Verlag.

Foster, P.J. & Burton, A. 2004. *Virtual Reality in Improving Mining Ergonomics. Journal of the South African Institution of Mining & Metallurgy,* 2003:129-133.

Foster, P.J. & Burton, A. 2006. *Modelling Potential Sightline Improvements to Underground Mining Vehicles Using Virtual Reality*. *Mining Technology*, 115(3):85-90.

Fowle, K. 2003. *The Use Of Computer Graphics And Visualisation (From Reconstruction To Training) For The Resource Sector Of Western Australia.* PhD thesis, University of Nottingham.

Gagne, R.M. & Merrill, M.D. 1990. Integrative goals for instructional design. *Educational Technology Research and Development*, 38(1):23-30.

Galimberti, C., Belloni, G., 2003. Three-dimensional virtual environments for cybertherapy: a psychosocial approach to effective usability. *Cyberpsychology & Behavior,* 6:229–236.

Gang, W., Jun, G. & Yingzhen, L. 2006. The Construction of Evaluation Framework for Virtual Geographic Environment. *Proceedings of the 16th International Conference on Artificial Reality and Telexistence*, ICART:20-25.

Genise, P. 2002. *Usability Evaluation: Methods and Techniques* [Online]. Available from: http://www.cs.utexas.edu/users/almstrum/cs370/elvisino/usaEval.html [Accessed: 15/03/09].

Geomagic.com. 2013. *Geomagic Haptic Devices* [Online]. Available from: http://geomagic.com/en/products-landing-pages/haptic [Accessed: 09/01/2014].

George, D. & Mallery, P. 2003. *SPSS for Windows step by step: A simple guide and reference*, 4th ed. Boston: Allyn & Bacon.

Georgsson, M, Weir, C.R. & Staggers, N. 2014. Revisiting Heuristic Evaluation Methods to Improve the Reliability of Findings. In: *C. Lovis (ed). e-Health – For Continuity of Care.* European Federation for Medical Informatics and IOS Press.

Gershon, N., Braham, R., Fracchia, D., Glassner, A., Mones-Hattal, B. & Rose, R. 1996. Breaking the myth: one picture is NOT (always) worth a thousand words. Panel transcript. *Proceedings of SIGGRAPH 1996,* 491.

Gillham, B. 2000. *Case Study Research Methods.* London: Bill Gillham.

Gliner, J.A. & Morgan, G.A. 2000. *Research Methods in Applied Settings: An Integrated Approach to Design and Analysis*. Mahwah, N.J: Lawrence Erlbaum.

Goldstone, R.L. & Son, J.Y. 2005. The transfer of scientific principles using concrete and idealized simulations. *Journal of the Learning Sciences,* 14(1):69-110.

Gomez, K. 2012. *BIM research will benefit energy and mining industries* [Online]. Available from: http://www.pacetoday.com.au/news/bim-research-will-benefit-energy-and-mining-indust [Accessed: 15/02/2015].

Gonzalez, R.A. & Sol H.G. 2012. Validation and Design Science Research in Information Systems. In: *M. Mora, O. Gelman, A. Steenkamp & M.S. Raisinghani (eds). Research Methodologies, Innovations and Philosophies in Software Systems Engineering and Information Systems:* 403-426. IGI Global.

Göransson, B. 2004. *User-Centred Systems Design: Designing Usable Interactive Systems in Practice*. Uppsala: Acta Universitatis Upsaliensis [Online]. Available from: http://uu.diva-portal.org/smash/get/diva2:164639/FULLTEXT01 [Accessed: 29/05/2010].

Govindasamy, T. 2002. Successful Implementation of e-Learning Pedagogical Considerations. *The Internet and Higher Education*, 4(4):287-299.

Gregor, S. 2002, Design Theory in Information Systems. *Australasian Journal of Information Systems*, Special Issue(2002):14-22.

Gregor, S. & Hevner, A.R. 2013. Positioning and Presenting Design Science Research for Maximum Impact. *MIS Quarterly*, 37(2):337–355. June 2013.

Guimarães, M.D.P. & Martins, V.F. 2014. A Checklist to Evaluate Augmented Reality Applications. *Proceedings of IEEE XVI Symposium on Virtual and Augmented Reality*, 2014:45-52.

Guion, L.A., Diehl, D.C. & McDonald, D. 2011. *Triangulation: Establishing the Validity of Qualitative Studies* [Online]. Available from: http://edis.ifas.ufl.edu/fy394 [Accessed: 21/05/2014].

Gulati, A. & Dubey, S.K. 2012. Critical Analysis on Usability Evaluation Techniques. *International Journal of Engineering Science and Technology,* 4(3):990–997.

Gulliksen, J., Goransson, B., Boivie, I., Blomkvist, S., Persson, J. & Cajander, A. 2003. Key Principles for user-centred Design. *Behaviour & Information Technology*, 22(6):397-409.

Hagan, T.O. 2004. Technology transfer of strata control knowledge. *SIMRAC Report* [Online]. Available from: www.mhsc.org.za/dmdocuments/Booklets/SIM020207.pdf [Accessed: 29/05/2010].

Hamilton-Atwell, A., Du Toit, A., Kirstein, L., Louw, L., Mtombeni, L. & Moses, R. 1997. Identification of Causes of Unsafe Acts or Neglect Resulting in Roof or Sidewall Accidents. *SIMRAC Report COL 307.*

Hannafin, M. J. 1992. Emerging technologies, ISD, and learning environments: Critical perspectives. *Educational Technology Research and Development*, 40(1):49-63.

Hanna, N. & Richards, D. 2014. Evaluation Framework for 3D Collaborative Virtual Environments (The CORE). *Proceedings of Pacific Asia Conference on Information Systems.* Paper 284.

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.

Hanssen, C.E., Lawrenz, F. & Dunet, D.O. 2008. Concurrent Meta-Evaluation, A Critique. *American Journal of Evaluation*, December 2008, 29(4): 572-582.

Harris, J. & Henderson, A. 1999. A Better Mythology for System Design. In: *M.W. Altom & M.G. Williams (eds). Proceedings of the ACM CHI 99 Human Factors in Computing Systems Conference,* May 15-20, 1999, Pittsburgh, PA:88-95.

Hay, K.E., Kim, B. & Roy, T.C. 2005. Design-based Research: More than Formative Assessment? An Account of the Virtual Solar System Project. *Educational Technology*. Jan-Feb 2005.

Hede, T. & Hede, A. 2002. Multimedia Effects on Learning: Design Implications of an Integrated Model. In: *S. McNamara & E. Stacey (eds). Untangling the Web: Establishing Learning Links, Proceedings ASET Conference 2002*, Melbourne July 7–10.

Heeter, C. 2000. Interactivity in the Context of Designed Experiences. *Journal of Interactive Advertising*, 1(1) [Online]. Available from:

http://jiad.org/vol1/no1/heeter/index.html [Accessed: 23/03/2009].

Henry, S.L. 2007. Just Ask: Integrating Accessibility Throughout Design [Online]. Available from: http://www.uiaccess.com/JustAsk/ [Accessed: 25/03/2009].

Herrington, J., McKenney, S., Reeves, T. & Oliver, R. 2007. Design-based research and doctoral students: Guidelines for preparing a dissertation proposal. In*: C. Montgomerie & J. Seale (eds). Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, 2007:4089-4097. Chesapeake, VA: AACE.

Hesham, A. 2004. Individual Differences: Implications for Web-based Learning Design. *International Education Journal,* 4(4):188–196.

Hevner, A.R. & Chatterjee, S. 2010. *Design Research in Information Systems*, Theory and Practice. New York: Springer.

Hevner, A.R., March, S.T., Park, J. & Ram, S. 2004. Design science in Information Systems Research. *MIS Quarterly,* 28(1):75-105.

Hix, D. & Gabbard, J.L. 2002. Usability Engineering of Virtual Environments. In: *K. Stanney (ed)*. *Handbook of Virtual Environments: Design, Implementation and Applications*. 681-699. Lawrence Erlbaum.

Hoadley, C.M. 2004. Methodological Alignment in Design-Based Research, *Educational Psychologist*, 39(4):203-212.

Hochstein, L. 2002. *GOMS and Keystroke-Level Model* [Online]. Available from: http://www.cs.umd.edu/class/fall2002/cmsc838s/tichi/printer/goms.html [Accessed: 15/03/2009].

Hogue, R. 2013. *Epistemological Foundations of Design-based Research* [Online]. Available from: http://www.slideshare.net/rhogue1/r-hogue-cssedbrv1 [Accessed: 19/10/2014].

Hollingsed, T. & Novick, D.G. 2007. Usability inspection methods after 15 years of research and practice. *SIGDOC '07 Proceedings of the 25th annual ACM international conference on Design of communication*, 249-255.

Hollender, N., Hofmann, C., Deneke, M. & Schmitz, B. 2010. Integrating cognitive load theory and concepts of human–computer interaction. *Computers in Human Behavior*, 26(2010):1278–1288.

Homan, D. 2010. *NASA Virtual Reality Lab Manager* [Online]. Available from: http://www.nasa.gov/audience/foreducators/spacesuits/careercorner/david-homan.html [Accessed: 22/03/2010].

Hörold, S., Mayas, C. & Krömker, H. 2012. Identifying the Information Needs of Users in Public Transport, In: *N. Stanton (ed). Advances In Human Aspects of Road and Rail Transportation,* 331-340. Boca Raton: CRC Press.

Howe, M. 2014. BIM Upgrades Australia's Mining Efficiency [Online]. Available from: http://sourceable.net/bim-upgrades-australias-mining-efficiency [Accessed: 15/02/2015].

Hvannberg, E.T., Halldórsdóttir, G. & Rudinsky, J. 2012. Exploitation of Heuristics for Virtual Environments. *Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*, 2012:308-317.

IITSEC (Interservice/Industry Training, Simulation and Education Conference). 2014. 'Squad Overmatch' Integrates Critical Army Skill Sets. *IITSEC Daily December 1, 2014* [Online]. Available from: http://www.iitsec.org/Documents/IITSEC_2014_Show_Daily/IITSEC14_Day1_web.pdf [Accessed: 11/01/2015].

Imenda, S. 2014. Is There a Conceptual Difference between Theoretical and Conceptual Frameworks? *Journal of Social Sciences*, 38(2):185-195.

Implats. 2010a. *Implats – Our Market* [Online]. Available from: http://www.implats.co.za/m/m_i.asp [Accessed: 25/05/2010].

Implats. 2010b: *Implats Annual Report* [Online]. Available from: http://www.implats.co.za/im/files/ar/2006/review06/impala.htm [Accessed: 25/05/2010].

Implats. 2010c. *Implats – Our Business* [Online]. Available from: http://www.implats.co.za/business/impala.asp [Accessed: 25/05/2010].

INUSE. 2004. *INUSE 6.2 Handbook of User-centred Design* [Online]. Available from: www.ejeisa.com/nectar/inuse/6.2/3-1.htm [Accessed: 17/06/2008].

International Mining, 2012. Mining's Virtual Reality. *International Mining Journal*, October 2012:16–27.

Iovine, J. 1995. *Step into Virtual Reality*. New York: Windcrest McGraw-Hill Inc.

ISO (International Organisation for Standardisation). 1998. *ISO 9241: Guidance on Usability Standards* [Online]. Available from: http://www.iso.ch/iso/en/CatalogueListPage.CatalogueList?ICS1=13&ICS2=180 [Accessed: 13/06/2008].

ISO. 2008. International Organisation for Standardisation ISO DIS 9241-210: Ergonomics of Human-System Interaction. *Part 210: Human-centred design process for interactive systems*. International Organisation for Standardisation. Geneva.

ISO 9241-11. 2008. ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDT's). Part 11 - Guidelines for specifying and measuring usability. Geneva: International Organisation for Standardisation.

ISO 9241-210. 2010. ISO 9241-210: Ergonomics of human-system interaction. Geneva: International Organisation for Standardisation.

Jamasmie, C. 2009. Mining in Five Dimensions. *Mining.com, January 2009* [Online]. Available from: http://magazine.mining.com/Issues/0901/MiningFiveDimensions.pdf [Accessed: 13/04/2010].

Jantke, K.P. & Knauf, R. 2005. Didactic design through storyboarding: standard concepts for standard tools. *WISICT '05 Proceedings of the 4th International Symposium on Information and Communication Technologies*, Cape Town, South Africa. January 2005.

Jentsch, F. & Bowers, C.A. 1998. Evidence for the validity of PC-based simulations in studying aircrew coordination. *The International Journal of Aviation Psychology,* 8(3):243-260.

Ji-zu, L., Li-mei, A., Jin-yun, W. & Xiao-li, Y. 2009. Research on simulation of safety behavior in coal mines based on the 3D and VIRTOOLS technologies. *ICCSIT, 2009 2nd IEEE International Conference on Computer Science and Information Technology*, 477-480.

John Deere. 2013. *John Deere reveals 2014 8R Tractor Using Virtual Reality* [Online]. Available from: http://www.savvyfarmgirl.com/2013/08/john-deere-reveals-2014-8r-tractor.html [Accessed: 13/01/2015].

Jonassen, D.H. 1994. Thinking Technology. *Educational Technology*, 34(4):34-37.

Jonassen, D.H. 1999. Designing constructivist learning environment. In: *C.M. Reigeluth (ed). Instructional-design theories and models: A new paradigm of instructional theory,* 2:215-239. Mahwah, NJ: Lawrence Erlbaum.

Jones, A., Scanlon, E., Tosunoglu, C., Morris, E., Ross, S., Butcher, P. & Greenberg, J. 1999. Contexts for evaluating educational software. *Interacting with Computers*, 11(5):499-516.

Jooste, C., Van Biljon, J. & Mentz, J. 2014. Usability evaluation for Business Intelligence applications: A user support perspective. *SA Computer Journal*, 53:32–44.

Joseph, D. 2004. The Practice of Design-Based Research: Uncovering the Interplay between Design, Research, and the Real-World Context. *Educational Psychologist*, 39(4):235-242.

Joubert, R. 2009. Avatar. *S A Computer Magazine*. 17(8):70-72.

Joung, Y.K. & Noh, S.D. 2014. Integrated modeling and simulation with in-line motion captures for automated ergonomic analysis in product lifecycle management. *Concurrent Engineering*, September 2014, 22(3):218-233.

Juuti, K. & Lavonen, J. 2006. Design-based Research in Science Education: One Step Towards Methodology. *Nordic Studies in Science Education*, 2006(4).

Kaiser, P.K., Henning, J.G., Cotesta, L. & Dasys, A. 2002. Innovations in Mine Planning and Design utilizing Collaborative Immersive Virtual Reality. *Proceedings of 104th CIM Annual General Meeting.* April 2002. Vancouver, BC.

Kalawsky, R. S. 1997. *Exploiting virtual reality techniques in education* [Online]. Available from: http://www.man.ac.uk/MVC//SIMA/vrtech/intro1.html [Accessed: 01/04/2007].

Kalawsky, R.S. 1999. VRUSE – A Computerised Diagnostic Tool for Usability Evaluation of virtual/synthetic environment systems. *Applied Ergonomics*, 30:11–25.

Kalawsky, R.S. 2000. *Human Factors Aspects of Virtual Design Environments in Education, JTAP Project 305* [Online]. Available from: www.jisc.ac.uk/uploaded_documents/jtap-048.pdf [Accessed: 12/08/2009].

Kalawsky, R. S. 2006. Exploiting Virtual Reality techniques in Education and Training. *Advanced VR Centre, Loughborough University of Technology* [Online]. Available from: http://www.agocg.ac.uk/reports/virtual/vrtech/title.htm [Accessed: 12/11/2009].

Kamppuri, M., Tedre, M. & Tukiainen, M. 2006. Towards the Sixth Level in Interface Design: Understanding Culture. Proceedings of the 5th ACM Conference on Human Computer Interaction in Southern Africa. Cape Town, South Africa, January 25-27. CHI-SA 2006:69-74.

Karat, J. 1997. Evolving the scope of user-centered design. *Communications of the ACM*, 40(7):33–38.

Karoulis, A. & Pombortsis, A. 2003. Heuristic Evaluation of Web-Based ODL Programs. In: *C. Ghaoui (ed). Usability Evaluation of Online Learning Programs*. Hershey, PA: Information Science Publishing.

Kieras, D. 2003. Model Based Evaluations. In: *J.A. Jacko & A. Sears (eds). The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*. Mahwah, N.J: Lawrence Erlbaum.

Kim, J.H., Park, S.H., Lee, H., Yuk, K.C. & Lee, H. 2001. Virtual Reality Simulations in Physics Education. *Paper presented at the Ed Media 2001 Conference*, June 25-30, 2001, Tampere, Finland.

Kinshuk, K., Lin T. & Patel A. 2008. User Adaptation in Supporting Exploration Tasks in Virtual Learning Environments. In: *J. Weiss, J. Nolan, J. Hunsinger & P. Trifonas (eds).*

The International Handbook of Virtual Learning Environments, Dordrecht, The Netherlands: Springer.

Kizil, M.S. 2003. Virtual Reality Applications in the Australian Minerals Industry. *31st International Symposium on Computer Applications in the Minerals Industries Held under the auspices of The South African Institute of Mining and Metallurgy*. 14-16 May 2003. Cape Town, South Africa.

Kizil, M.S. & Joy J. 2001. What can Virtual Reality do for Safety? Queensland Mining Industry Health and Safety Conference: Managing Safety to Have a Future. *Queensland Mining Council Publication*, 173-181.

Kizil, M.S., Kerridge, A & Hancock, M. 2004. Use of Virtual Reality in Mining Education and Training. *Proceedings of the 2004 CRCMining Research and Effective Technology Transfer Conference*, 15-16 June.

Kjeldskov, J., Skov, M.B. & Stage, J. 2004. Instant Data Analysis: Conducting Usability Evaluations in a Day. *Proceedings of the third Nordic conference on Human-Computer Interaction,* 2004:233-240. Tampere: ACM International Conference Proceeding Series.

Kline, P. 2000. *The Handbook of Psychological Testing*, 2nd ed. London: Routledge.

Kneebone, R., Arora, S., King, D., Bello, F., Sevdalis, N., Kassab, E., Aggarwal, R., Darzi, A. & Nestel, D. 2010. Distributed simulation – Accessible immersive training. *Medical Teacher,* 32(1):65-70.

Kowalski-Trakofler, K.M. & Barrett, E.A. 2003. The concept of degraded images applied to hazard recognition training in mining for reduction of lost-time injuries. *Journal of Safety Research*, 34:515-525.

Krige, B. 1995. Leon Commission: A dismal record. *SA Mining, Coal, Gold & Base Minerals*, May 1995. Johannesburg, 9-13.

Labaree, R.V. 2015. *Organizing Your Social Sciences Research Paper* [Online]. Available from: http://libguides.usc.edu/writingguide [Accessed 21/02/2015].

Lartigue, L., Cathcart, J., Kelleher, T.M., Pfundstein, A. & Williams. M.D. 2014. Leveraging Oculus Rift for an Immersive Distance-Learning Experience: A High

Definition, Panoramic Lecture Recording/Playback System using Commercial Virtual Reality Tools. *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, 2014(1):1090–1099.

Lartigue, L., Scoville, T. & Pham, M. 2014. Promoting K-8 Learning using Oculus Rift: Employing Virtual Reality to Increase Learning Outcomes in Elementary Biology. *World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education,* 2014(1):1100–1105.

Lau, R.W.H., Yen, N.Y., Li, F. & Wah, B. 2014. Recent development in multimedia e-learning technologies. *World Wide Web Journal,* 17(2):189-198. Springer.

Lazar, J., Feng, J.H. & Hochheiser, H. 2010. *Research Methods in Human-Computer Interaction*. UK: John Wiley & Sons.

Lee, A.T. 2005. *Flight simulation: virtual environments in aviation*. Hampshire, UK: Ashgate.

Lennie, J., Tacchi, J. & Wilmore, M. 2012. Meta-Evaluation to improve learning, evaluation capacity development and sustainability: findings from a participatory evaluation project in Nepal. *South Asian Journal of Evaluation in Practice*, 1(1):13-28.

Le Roux, A. 2005. Mine fatalities down, but still too high. *Mining Weekly* [Online]. Available from: http://www.miningweekly.co.za/min/features/health/?show=66581 [Accessed: 12/06/2007].

Leshem, S. & Trafford, V. 2007. Overlooking the conceptual framework. *Innovations in Education and Teaching International,* 44(1):93-105.

Levi, M.D. & Conrad, F.G. 1996. A Heuristic Evaluation of a World Wide Web Prototype. *Interactions*, 3(4):50-61.

Levy, J.R. & Bjelland, H. 1995. *Create your own virtual reality system*. New York: Windcrest McGraw-Hill Inc.

Lewis, C.H. & Griffin, M.J. 1997. Human factors consideration in clinical applications of virtual reality. In: *G.Riva (ed). Virtual Reality in Neuro-Psycho-Physiology. Cognitive,* *Clinical and Methodological Issues in Assessment and Rehabilitation,* Italy, IOS Press, 35-56.

Li, H.M. & Kang, B.S. 2014. A Virtual Reality Safety Training System for Coal Mining Industry. *Applied Mechanics and Materials*, 687:2371-2374.

Liao, Y. 1999. Effects of Hypermedia on Student Achievement: A meta-analysis. *Journal of Educational Multimedia and Hypermedia*, 8(3):255-277.

Liarokapis, F., Sylaiou, S., Basu, A., Mourkoussis, M., White, M. & Lister, P.F. 2004. An Interactive Visualisation Interface for Virtual Museums. *The 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage.* Oudenaarde, Belgium, 1-10.

Lourens, C. 2009. *Impala Platinum's Worst Ever Mining Accident Kills 9* [Online]. Available from: http://www.bloomberg.com/apps/news?pid=20601116&sid=aOa71EYXvJSQ [Accessed: 31/05/2010].

Lubbe, O.S. 2006. Interview with Chief Executive, Chrome Mining Division, Xstrata Alloys, Rustenburg. (Notes in possession of author).

Lucas, J. & Thabet, W. 2008. Implementation and Evaluation of a VR Task-based Training Tool for Conveyor Belt Safety Training. *ITCON Special Issue Virtual and Augmented Reality in Design and Construction*, 13:637-659.

Lucas, J., Thabet, W. & Worlikar, P. 2007. Using Virtual Reality (VR) to Improve Conveyor Belt Safety in Surface Mining. *ITC Digital Library* [Online]. Available from: http://itc.scix.net/cgi-bin/works/Show?w78_2007_42 [Accessed: 10/06/2009].

Ma, Y. & Harmon, S.W. 2009. A Case Study of Design-Based Research for Creating a Vision Prototype of a Technology-Based Innovative Learning Environment. *Journal of Interactive Learning Research,* 20(1):75 -93.

MacDonald, R.J. 2008. Professional Development for Information Communication Technology Integration: Identifying and Supporting a Community of Practice through Design-Based Research. *Journal of Research on Technology in Education,* 40(4):429-445.

Madan, A. & Dubey, S.K. 2012. Usability Evaluation Methods: a literature review. *International Journal of Engineering Science and Technology*. 4(2):590–599.

Magner, U.I.E., Schwonke, R., Aleven, V., Popescu, O. & Renkl, A. 2013. Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction*, 29:141-152.

Maguire, M. 2001. Context of Use within usability activities. *International Journal of Human-Computer Studies,* 55:453-483.

Manic, L.J., Aleksic, M. & Tankosic, M. 2013. Possibilities of New Technologies in Promotion of the Cultural Heritage: Danube Virtual Museum, WSEAS, Advances in Environment, Ecosystems and Sustainable Tourism, *Proceedings of the 2nd International Conference on Sustainable Tourism and Cultural Heritage (STACH '13)*, Brasov, Romania, June 1-3, 322-326.

Mao, J., Vredenburg, K., Smith, P.W. & Carey, T. 2005. The State of User-centred Design Practice*. Communications of the ACM*, Mar2005, 48(3):105-109.

March, S.T. & Smith, G. 1995. Design and Natural Science Research on Information Technology. *Decision Support Systems,* 15(4):251-266.

March, S.T. & Storey, V.C. 2008. Design Science in the Information Systems Discipline: An introduction to the special issue on design science research, *MIS Quarterly*, 32(4):725–730.

Marchand, A. & Walker, S. 2009. Designing in Design Research: From Solving Problems to Exploring Issues, *Proceedings of the 8th European Academy Of Design Conference*, The Robert Gordon University, Aberdeen, Scotland.

Markus, M.L., Majchrzak, A. & Gasser, L. 2002. A Design Theory for Systems that support emergent Knowledge Processes. *MIS Quarterly,* 26(3):179-212.

Marsh, T., Wright, P. & Smith, S. 2001. Evaluation for the design of experience in virtual environments: modelling breakdown of interaction and illusion. *Cyberpsychology & Behavior,* 4(2):225–238.

Masemola, S.S. & De Villiers, M.R. 2006. Towards a Framework for Usability Testing of Interactive e-Learning Applications in Cognitive Domains, Illustrated by a Case Study. In: *J. Bishop & D. Kourie (eds). Service-Oriented Software and Systems. Proceedings of SAICSIT 2006*: 187-197. ACM International Conference Proceedings Series.

Mason, R., Cooper, G. & Wilks, B. 2015. Using Cognitive Load Theory to select an Environment for Teaching Mobile Apps Development. *Proceedings of Australasian Computing Education Conference ACE2015.* Sydney, Australia. January 2015.

Matthee, M., Henneke, M. & Johnson, R. 2014. Tensions in the adoption of e-Learning in the mining industry of South Africa. *South African Computer Journal,* 52:42–54.

Mayer, R.E. 2002. *Multimedia Learning*. New York: Cambridge University Press.

Mayer, R.E. 2008. Applying the Science of Learning: Evidence-based Principles for the Design of Multimedia Instruction. *American Psychologist,* 63(8):760–769.

Mayer, R.E. 2014. Incorporating motivation into multimedia learning. *Learning and Instruction,* 29:171–173.

Mayer, R.E., Fennell, S., Farmer, L. & Campbell, J. 2004. A Personalization Effect in Multimedia Learning: Students Learn Better when Words are in Conversational Style rather than Formal Style. *Journal of Educational Psychology,* 96(2):389–395.

Mayer, R.E. & Moreno, R. 2003. Nine ways to reduce cognitive load in multimedia learning. *Educational psychologist*, 38(1):43–52.

Mayes, T. & De Freitas, S. 2005. *JISC e-Learning Models Desk Study Stage 2: Review of e-learning theories, frameworks and models [Online]. Available from:* http://www.uwic.ac.uk/ltdu/docs/Review%20of%20e-learning%20theories.pdf *[Accessed: 13/04/2010].*

Mayes, J.T. & Fowler, C.J. 1999. Learning Technology and Usability: A Framework for Understanding Courseware. *Interacting with Computers*, 11(5):485-497.

McCombs, B.L. & Whisler, J.S. 1997. *The learner-centered classroom and school: Strategies for increasing student motivation and achievement.* San Francisco: Jossey-Bass.

McCombs, B.L. & Vakili, D. 2005. A Learner-Centered Framework for E-Learning. *Teachers College Record,* 107(8):1582-1600.Columbia University.

McGaghie, W.C. 2007. Advances in Simulation-based Medical Education and Research. *2007 Jack Maatsch Address at Feinberg School of Medicine, Northwestern University* [Online]. Available from: http://omerad.msu.edu/endowments/WilliamMcGaghie_Lecture.pdf. [Accessed: 08/04/2010].

McKay, D. 2009. *Law and Justice* [Online]. Available from: http://www.klasslooch.com/91912.html [Accessed: 05/06/2010].

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

McNiff, J. 2013. *Action Research: Principles and Practice.* 3rd ed. New York: Routledge.

Metzger, M.J., Flanagin, A.J. & Medders, R.B. 2010. Social and Heuristic Approaches to Credibility Evaluation Online. *Journal of Communication,* 60(3):413–439.

Mereu, S.W. & Kazman, R. 1996. Audio enhanced 3D interfaces for visually impaired users. *Proceedings of CHI '96 Conference:*72-78, Vancouver, BC.
MHSC. 2010. *Mine Health and Safety Council* [Online]. Available from: www.mhsc.org.za [Accessed: 01/06/2010].

MHSA (Mine Health and Safety Act). 1996. *Act No. 29 of 1996* [Online]. Available from: http://www.polity.org.za/html/govdocs/legislation/1996/act96-029.html [Accessed: 22/04/2007].

Miles, M.B. & Huberman, A.M. 1994. *An expanded sourcebook: Qualitative Data Analysis*. California, USA: SAGE Publications.

Milgram, P. & Kishino, F. 1994. A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information Systems*, E77-D(12).

Mills, S. 2001. The importance of task analysis in Usability Context Analysis – designing for fitness for purpose. *Behaviour & Information Technology*, 19(1):57–68.

Mills, S. 2007. Contextualising Design: Aspects of using usability context analysis and hierarchical task analysis for software design. *Behaviour & Information Technology*, 26(6):499–506.

Mingming, G. & Keping, Z. 2009. Virtual Reality Simulation System for Underground Mining Process. *Proceedings of IEEE 2009 4th International Conference on Computer Science & Education*. Nanning, China.

Mining Weekly. 2002. *Mine health and safety on the road to recovery* [Online]. Available from: http://www.miningweekly.co.za/min/features/health/?show=21234.
Published: 2002/05/10 [Accessed: 12/06/2007].

Mining Weekly. 2007. *SA to take on mining industry over safety record* [Online].
Available from: http://www.miningweekly.co.za/article.php?a_id=111499
Published: 2007/06/22 [Accessed: 05/07/2007].

Misu, T., Georgila, K., Leuski, A. & Traum, D. 2012. Reinforcement learning of question-answering dialogue policies for virtual museum guides. *Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue,* 84-93. Association for Computational Linguistics.

Mitra, R. & Saydam, S. 2013. Using Virtual Reality in Tertiary Education. *International Journal of Technologies in Learning*. 19:97–112.

Moczarny, I.M., De Villiers, M.R. & Van Biljon, J.A. 2012. How can usability contribute to user experience? A study in the domain in e-commerce. *SAICSIT '12*, Pretoria, South Africa: ACM, 216-225.

Moldenhauer, R. 2006. Interview with Safety, Health and Environment Manager, Anglo Platinum Lebowa. (Notes in possession of author).

Molenda, M. 2003. In search of the elusive ADDIE model. *Performance Improvement*, 42(5):34-37.

Molich, R. & Nielsen, J. 1990. Improving a Human-Computer Dialogue. *Communications of the ACM*, 33(3):338-348.

Monforton, C. & Windsor, R. 2010. An Impact Evaluation of a Federal Mine Safety Training Regulation on Injury Rates Among US Stone, Sand, and Gravel Mine Workers: An Interrupted Time-Series Analysis. *American Journal of Public Health,* July 2010, 100(7):1334–1341.

Moreno, R. & Mayer, R.E. 1999. Cognitive principles of multimedia learning: the role of modality and contiguity. *Journal of Educational Psychology*, 91(2):358–368.

Moreno, R. & Mayer, R.E. 2000. A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages. *Journal of Educational Psychology*, 97(1):117–125.

Moreno, R. & Mayer, R.E. 2002. Learning science in virtual reality multimedia environments: role of methods and media. *Journal of Educational Psychology*, 94(3):598-610.

Morrison, B.B., Dorn, B. & Guzdial, M. 2014. Measuring cognitive load in introductory CS: adaptation of an instrument. *Tenth International Conference on Computing Education Research (ICER2014)*, Glasgow, Scotland, 131-138.

Mosqueira-Rey, E., Alonso-Rios, D. & Moret-Bonillo, V. 2009. Usability Taxonomy and Context-of-Use Taxonomy for Usability Analysis. *Proceedings of the 2009 IEEE International Conference on Systems, Man and Cybernetics*, San Antonio, Texas, USA.

MQA. 2014. *Mining Qualifications Authority* [Online]. Available from: www.mqa.org.za [Accessed: 14/11/2014].

MSHA, 1998. Mine Safety and Health Administration: Educational Field Services. *A New Focus in MSHA* [Online], Available from: www.msha.gov/training/epdfs/epdfs.pdf [Accessed: 10/01/2006].

Munoz, R. & Chalegre, V. 2012. Defining Virtual Worlds Usability Heuristics. *Proceedings of Ninth International Conference on Information Technology: New Generations (ITNG)*. Las Vegas, 2012:690-695.

Nalbant, G. & Bostan, B. 2006. Interaction in Virtual Reality. *Proceedings of 2006 International Symposium of Interactive Media Design* [Online]. Available from: http://newmedia.yeditepe.edu.tr/pdfs/isimd_06/12.pdf [Accessed: 10/02/2010].

Nasios, K. 2001. Improving Chemical Plant Safety Training Using Virtual Reality. *PhD Thesis*, University of Nottingham.

NATO (North Atlantic Treaty Organisation), 2003. Virtual Reality: State of Military Research and Applications in Member Countries. *Technical Report from the Research and Technology Organisation of the North Atlantic Treaty Organisation* [Online]. Available from: http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA411978 [Accessed: 06/04/2010].

Ndaba, D. 2008. Can South Africa stop the mine fatalities? *Mining Weekly* [Online]. Available from: http://www.miningweekly.com/article/can-south-africa-stop-the-mine-fatalities-2008-02-01 [Accessed: 05/06/2010].

News24. 2007. *News24* [Online]. Available from: http://www.news24.com/News24/South_Africa/Politics/0,,2-7-12_2196310,00.html [Accessed: 05/10/2007].

Nielsen, J. 1994. Heuristic Evaluations. In: *J. Nielsen & R.L. Mack (eds). Usability Inspection Methods*. New York: John Wiley & Sons.

Nielsen, J. 2005. *Ten Usability Heuristics* [Online]. Available from: www.useit.com/papers/heuristic/heuristic_list.html [Accessed: 04/12/2014].

Noffke, S. & Somekh, B. 2009. *Handbook of Educational Action Research*. London: Sage.

Noor, A.K. & Wasfy, T.M. 2001. Simulation of Physical Experiments in Immersive Virtual Environments. *Engineering Computations*, 18(3):515-538.

Norman, D.A. & Draper, S.W. 1986. *User Centred Systems Design.* Hillsdale, NJ: Lawrence Erlbaum.

Norman, D.A, & Spohrer, J.C. 1996. Learner-Centered Education. *Communications of the ACM,* 39(4):26-27.

Notess, M. 2001. Usability, User experience, and Learner Experience. *eLearn Magazine, ACM Digital Library* [Online]. Available from: http://www.elearnmag.org [Accessed: 12/12/2013].

NPC. 2013. *National Planning Commission - Mineral resources* [Online]. Available from: http://www.npconline.co.za/pebble.asp?relid=163 [Accessed: 14/05/2014].

Nyang'or, J.O., De Villiers, M.R. & Ssemugabi, S. 2013. A framework for usability evaluation of an offline e-learning tutorial and its application in usability testing. *World Conference on Educational Multimedia, Hypermedia and Telecommunications (EDMEDIA)* 2013. Victoria, Canada.

Oates, B. 2006. *Researching Information Systems and Computing*. London: Sage Publications.

Oculus, 2015. *Oculus VR* [Online]. Available: http://www.oculusvr.com [Accessed: 12/01/2015].

Odendaal, N. 2014. *Harmony accident major setback – Shabangu* [Online]. Available from: http://www.miningweekly.com/article/harmony-accident-major-setback-shabangu-2014-02-13 [Accessed: 12/10/2014].

Olivier, M.S. 2004. *Information Technology Research: A Practical Guide for Computer Science and Informatics*. 2nd ed. Van Schaik, Pretoria.

Olsen, D.R. 2010. *Building Interactive Systems: Principles for Human-Computer Interaction*. Course Technology, Cengage Learning. Boston, MA.

Oosthuizen, B. 2009. *Realistic Simulation Accelerates Safety Evaluation of Mine Designs* [Online]. Available from: http://magazine.mining.com/Issues/0901/SimulationSafetyEvaluationMineDesigns.pdf [Accessed: 03/04/2010].

Orr, T.J., Filigenzi, M.T. & Ruff, T.M. 2002. *Desktop Virtual Reality Miner Training Simulator* [Online]. Available from: www.cdc.gov/niosh/mining/pubs/pdfs/dvrmt.pdf [Accessed: 12/01/2008].

Orr, T.J., Mallet, L.G. & Margolis, K.A. 2009. Enhanced fire escape training for mine workers using virtual reality simulation. *Mining Engineering,* 61(11):41-44.

Oryx, 2014. *Oryx launches Volvo simulators* [Online]. Available from: http://www.miningmagazine.com/technology/simulation/ [Accessed: 10/01/2015].

Oshhima, T., Yamamoto, H. & Tamura, H., 1996. Gaze-directed adaptive rendering for interacting with virtual space. *Proceedings of 1996 IEEE Virtual Reality Annual International Symposium*:103–110, Santa Clara, California.

Osterlund, J. & Lawrence, B. 2012. Virtual reality: Avatars in human spaceflight training. *Acta Astronautica*, 71:139-150.

Ottosson, S. 2002. Virtual reality in the product development process. *Journal of Engineering Design*, 13(2):159–172.

Oviatt, S. 2006. Human-Centered Design Meets Cognitive Load Theory: Designing Interfaces that Help People Think. *Proceedings of the 14th annual ACM international conference on Multimedia*:871–880. New York.

Paas, F., Renkl, A. & Sweller, J. 2003. Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist,* 38(1):1-4.

Paddison, C. & Englefield, P. 2003. Applying heuristics to perform a rigorous accessibility inspection in a commercial context. In: *M. Zajicek & A. Edwards (eds). Proceedings of the 2003 conference on Universal usability,* 126-133. New York, NY: ACM Press.

Patel, H., Stefani, O., Sharples, S., Hoffmann, H., Karaseitanidi, I. & Amditis, A. 2006. Human centred design of 3-D interaction devices to control virtual environments. *International Journal of Human-Computer Studies*, 64(3):207-220.

Pimentel, K. & Teixeira, K. 1993. *Virtual reality: through the new looking glass*. New York: McGraw-Hill.

Pinelle, D., Wong, N. & Stach, T. 2008. Heuristic evaluation for games: usability principles for video game design. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York:1453-1462.

Pitzer, C.J. 2000. A Study of the Risky Positioning Behaviour of Operators of Remote Controlled Mining Equipment. *Special Report MDG 5004*, New South Wales: Department of Mineral Resources.

PMHSA, 2008. *Presidential Mine Health and Safety Audit Report* [Online]. Available from: http://www.dme.gov.za/pdfs/speeches/Presidential%20Audit/Presidential%20Audit%20 Report.pdf [Accessed: 17/05/2010].

Plomp, T. 2000. *Development Research in/on Educational Development* [Online]. Available from: http://www.cite.hku.hk/events/doc/2000/DevResHKGsumm0B20.doc [Accessed: 12/04/2008].

Plomp, T. 2002. *Some reflections on development research (DR)*. Seminar at the Faculty of Education, University of Pretoria. August 2002.

Plomp, T. 2007. Educational design research: An introduction. In: *T. Plomp & N. Nieveen (eds). An introduction to educational design research*. SLO: NICD.

Pollock, E., Chandler, P. & Sweller, J. 2002. Assimilating complex information. *Learning and Instruction,* 12*:*61-86.

Pool, A. & Ball, L.J. 2006. Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects. In: *C. Ghaoui (ed). Encyclopaedia of Human-Computer Interaction*. Idea Group. Pennsylvania:211–219.

Poppe, R., Rienks, R. & Van Dijk, B., 2007. *Evaluating the Future of HCI: Challenges for the Evaluation of Emerging Applications for Human Computing.* Lecture Notes in Artificial Intelligence, 234-250. Springer Verlag.

Powell, G.C. 2001. The ABC of Online Course Design*. Educational Technology,* 41(1):43-47.

Poynton, L. 2009. Evaluating the Effectiveness of Virtual Reality Learning in a Mining Context University of New South Wales, Australia. *Master's Thesis*.

Prantosh, P., Kalyan, K., Dipak, C. & Rajesh, R. 2014. Usability engineering and user interface design for electronic information systems and its subsystems. *Information Studies*, 20(1):23–32.

Preece, J. 1993. *A Guide to Usability: Human Factors in Computing*. The Open University: Addison-Wesley.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. & Carey, T. 1994. *Human-Computer Interaction.* Essex, England: Addison-Wesley Longman.

Preece, J., Rogers, Y. & Sharp, H. 2002. *Interaction Design: Beyond Human-Computer Interaction.* 1st ed. New York: John Wiley & Sons.

Preradović, N.M., Miličić, D. & Đuričić, P. 2014. A Model of Service Learning and Outreach for Primary Education through Museums. *International Journal of Education and Information Technologies,* 8:49-55.

Pressly, D. 2008. *New legislation to enforce health and safety rules in mines* [Online]. Available from: http://www.busrep.co.za/index.php?fArticleId=4468516 [Accessed: 05/06/2010].

Pringle, C. 2009. SA government says mine deaths decreased in 2008. *Mining Weekly*. [Online]. Available from: http://www.miningweekly.com/article/sa-government-says-mine-deaths-decreased-in-2008-2009-01-07 [Accessed: 05/06/2010].

Ready, D.J., Pollack, S., Rothbaum, B.O. & Alarcon, R.D. 2005. Virtual Reality Exposure for Veterans with Posttraumatic Stress Disorder. *Journal of Aggression, Maltreatment & Trauma*, 12(1):199–220.

Rebelo, F. & Noriega, P. 2012. Using Virtual Reality to Assess User Experience. *Human Factors*, 54(6):964-982.

Reeves, T.C. 1994. Evaluating what really matters in Computer-based Education. In: *M. Wild & D. Kirkpatrick (eds). Computer Education: New Perspectives*. Edith Cowan University Press. Perth, Australia, 219-246.

Reeves, T.C. 2000. Enhancing the worth of instructional technology research through 'design experiments' and other development research strategies. *Symposium on International perspectives on instructional technology research for the 21st century*. New Orleans, LA, USA.

Reeves, T. C. 2006. Design research from a technology perspective. In: *J.V.D. Akker, K. Gravemeijer, S. McKenney & N. Nieveen (eds). Educational design research.* New York: Routledge:52-66.

Reeves, T.C., Benson. L., Elliot, D., Grant, M., Holschuh, D., Kim, B., Kim, H., Lauber, E. & Loh, S. 2002. Usability and Instructional Design Heuristics for E-Learning Evaluation. *Proceedings of EDMEDIA 2002 World Conference on Educational Multimedia, Hypermedia and Telecommunications*, Denver, June 2002.

Reeves, T.C., Herrington, J. & Oliver, R. 2004. A development research agenda for online collaborative learning. *Educational Technology Research and Development,* 52(4):53–63.

Reeves, T.C., Herrington, J. & Oliver, R. 2005. Design Research: A socially Responsible Approach to Instructional Technology Research in Higher Education. *Journal of Computing in Higher Education*, *16*(2):97-116.

Reeves, T.C. & Reeves, P.M. 1997. Effective Dimensions of Interactive Learning on the World Wide Web. In: *B.H. Khan (ed). Web-Based Instruction*. New Jersey: Educational Technology Publications.

Reeves, T.C., West, R. & Orrill, C. 2006. *EDIT 9990 online discussion on Design-based Research* [Online]. Available from: http://it.coe.uga.edu/~treeves/EDIT9990/index.html *[Accessed: 12/02/2011].*

Reigeluth, C.M. 2013. *Instructional Design Theories and Models Volume II: A New Paradigm of Instructional Theory.* London: Lawrence Erlbaum.

Reigeluth, C.M. & Moore, J. 1999. Cognitive Education and the Cognitive Domain. In: *C.M. Reigeluth (ed). Instructional-Design Theories and Models*. Manwah, NJ: Lawrence Erlbaum.

Ren, B., Kong, W. & Ren, G. 2012. The Design and Realization of 3D Training System of Virtual Reality Coal Mine (3DTS-VRCM) Based on Multi-role. *Advances in Biomedical Engineering*, 6:363-369.

Rheingold, H. 1991. *Virtual Reality*, New York: Summit Books.

Richard, P., Birebent, G., Coiffet, P., Burdea, G., Gomex, D. & Langrana, N. 1996. Effect of frame rate and force feedback on virtual manipulation. *Presence: Teleoperators and Virtual Environments*, 5(1):95–108.

Richey, R.C., Klein, J.D. & Nelson, W.A. 2003. Development research: Studies of instructional design and development. In: *D.H. Jonassen (ed). Handbook of Research for Educational Communications and Technology*:1099–1130. London: Macmillan.

Richey, R. C. & Nelson, W. A. 1996. Developmental research. In: *D.H. Jonassen (ed). Handbook of Research for Educational Communications and Technology*:1213–1245. London: Macmillan.

Ritchie, D.C. & Hoffman, B. 1997. Incorporating Instructional Design Principles with the World Wide Web. In: *B.H. Khan (ed). Web-Based Instruction*. New Jersey: Educational Technology Publications.

Ritter, N. 2010. Understanding a widely misunderstood statistic: Cronbach's alpha. Paper presented at *Southwestern Educational Research Association (SERA) Conference* 2010: New Orleans, LA.

Rizzo, A.A., Klimchuk, D., Mitura, R., Bowerly, T., Buckwalter, J.G. & Parsons, T. 2006. A Virtual Reality Scenario for All Seasons: The Virtual Classroom. *CNS Spectrums,* 11(1):35-44.

Roberts, D. & Warwick, K. 1993. An overview of virtual reality. In: *K. Warwick, J. Gray & D. Roberts (eds). Virtual Reality in Engineering*. The Institution of Electrical Engineers, London, 1-22.

Robles-De-La-Torre, G. 2009. Virtual Reality: Touch/Haptics. In: *B. Goldstein (ed). Encyclopedia of Perception,* 2:1036–1038. Sage Publications, Thousand Oaks, California.

Rogers, Y., Sharp, H. & Preece, J. 2011. *Interaction Design – beyond human-computer interaction*. 3rd ed. John Wiley & Sons Ltd, England.

Rossi, M. & Sein, M. 2003. *Design Research Workshop: A Proactive Research Approach.* Presentation delivered at IRIS 26, August 9–12, 2003 [Online]. Available from: http://tiesrv.hkkk.fi/iris26/presentation/workshop_designRes.pdf_[Accessed: 24/03/2011].

Rothbaum, B.O., Anderson, P., Zimand, E., Hodges, L., Lang, D. & Wilson, J. 2006. Virtual Reality Exposure Therapy and Standard (in Vivo) Exposure Therapy in the Treatment of Fear of Flying. *Behavior Therapy*, 37(1):80-90.

Ruffini, M.F. 2000. Systematic Planning in the Design of an Educational Web Site. *Educational Technology*, 40(2):58-64.

Rusu, C., Muñoz, R., Roncagliolo, S., Rudloff, S., Rusu, V. & Figueroa, A. 2011. Usability Heuristics for Virtual Worlds. *Proceedings of AFIN 2011: The Third International Conference on Advances in Future Internet*, 2011:16-19.

Sachs, P. 1995. Transforming work: collaboration, learning and design. *Communications of the ACM,* Sept. 1995, 38(9).

Sadowski, W. & Stanney, K.M. 2002. Measuring and managing presence in virtual environments. *Handbook of virtual environments: Design, Implementation, and Applications*:791–806*.* Mahwah, NJ: Lawrence Erlbaum.

Sanders, J.R. & JCSEE (The Joint Committee on Standards for Educational Evaluation). 1994. *The Program Evaluation Standards: 2nd ed. How to Assess Evaluations of Educational Programs*, Sage Publications.

Sandoval, W.A. 2004. Developing Learning Theory by Refining Conjectures Embodied in Educational Designs. *Educational Psychologist*, 39(4):213-223.

Sandoval, W.A. & Bell, P. 2004. Design-based research methods for studying learning in context: Introduction. *Educational Psychologist*, *39*(4):199-201.

Sangrà, A., Vlachopoulos, D., & Cabrera, N. 2012. Building an inclusive definition of e-learning: An approach to the conceptual framework. *The International Review of Research in Open and Distributed Learning,* 13(2):145-159.

Sarathay, S., Shujace, K. & Cannon, K. 2000. Visualization of Large Complex Datasets Using Virtual Reality. *Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'00)*, Las Vegas, 252.

Sarraipa, J., Gomes-de-Oliveira, P., Marques-Lucena, C., Jardim-Goncalves, R. & Mendonça da Silva, J. 2013. E-Training Development Approach for Enterprise Knowledge Evolution. *ASME 2013 International Mechanical Engineering Congress and Exposition. Volume 5: Education and Globalization.* San Diego, California, USA, November 15–21, 2013.

SAS, 2014. *SAS 9.2 Procedures Guide: Statistical Procedures*. 3rd ed. 2014 [Online]. Available from: https://support.sas.com/documentation/cdl/en/procstat/63104/HTML/default/viewer.htm#procstat_freq_sect032.htm [Accessed: 14/05/2014].

Sawicka, A. 2008. Dynamics of cognitive load theory: A model-based approach. *Computers in Human Behavior*, 24(2008):1041–1066.

Schafrik, S.J., Karmis, M. & Agioutantis, Z. 2003. Methodology of Incident Recreation Using Virtual Reality. *Proceedings of SME Annual Meeting February 2003*. Cincinnati, Ohio.

Schmid, M. 2003. Virtually Assured – the role of virtual reality in the field of coal mining. *World Coal*, 11:45-48.

Schmid, M. & Bracher, P. 2004. Use of Conceptual Models in Managing Process and Product Innovations, *Proceedings of MinEXPO.* Las Vegas, September 2004.

Schmid, M. & Rossman, M. 2004. Virtual Reality teaching and training systems in mining – applications and operational experiences. *Gluckauf Mining Reporter 2004*.

Schmid, M. & Winkler, T. 2008. Interactive 3D supports underground repair and maintenance of Road Heading Machines and Shearer Loaders. *Proceedings of the 21st World Mining Congress*. Krakow, Poland. September 2008.

Schofield, D. 2014. A virtual education: Guidelines for using games technology. *Journal of Information Technology Education: Innovations in Practice*, 13:25-43.

Schofield, D., Denby, B. & Hollands, R. 2001. Mine Safety in the Twenty-First Century: The Application of Computer Graphics and Virtual Reality, In: *M. Karmis (ed). Mine Health and Safety Management*, Littleton, CO:SME, Chapter 10:153-174.

Schofield, D. & Lester, E. 2010. Learning in a Virtual Environment: Should there be Guidelines? *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications,* 2010:3848-3860. Chesapeake, VA: Association for the Advancement of Computing in Education (AACE).

Schreiber, W. & Kielblock, J. 2004. Self-contained self-rescuer legislation within the context of the Mine Health and Safety Act of South Africa: A critical analysis. *Journal of the Mine Ventilation Society of South Africa*, 57(4):119-123.

Scienceclarified. 2009. *Virtual Reality in Art and Entertainment* [Online]. Available from: http://www.scienceclarified.com/scitech/Virtual-Reality/Arcades-to-Avatars-Virtual-Reality-in-Art-and-Entertainment.html [Accessed: 22/03/2010].

Scriven, M. 2008. Evaluating Evaluations: A Meta-Evaluation Checklist. In: *C. L. S. Coryn & M. Scriven (eds). Reforming the evaluation of research. New Directions for Evaluation,* 118:61–72.

Seccombe, A. 2007. *SA mine safety under fire* [Online]. Available from: http://www.miningmx.com/grnbk2007/676219.htm [Accessed: 06/06/2010].

Second Life, 2012. *What is an Avatar?* [Online]. Available from: http://secondlife.com/whatis/avatar/?lang=en-US [Accessed: 17/03/2012].

Seggie, E. 2010. Health and safety secrets shared in the industry. *Mining Weekly* [Online]. Available from: http://www.miningweekly.com/article/health-and-safety-secrets-shared-in-the-industry-2010-05-14/page:70 [Accessed: 05/06/2010].

Sener, J. 1997. Constructivism: Asynchronous Learning Networks. *ALN Magazine* 1(1) [Online]. Available from: http://www.aln.org/publications/magazine/v1n1/sener/sener.asp [Accessed: 13/06/2009].

Sguazzin, A. & Lourens, C. 2007. *Miners' strike puts safety over output for a day* [Online]. Available from: http://www.busrep.co.za/index.php?fSectionId=561&fArticleId=4158940 [Accessed: 14/05/2010].

Sharma, G., Mavroidis, C. & Ferreira, A. 2005. Virtual Reality and Haptics in Nano- and Bionanotechnology. *Handbook of Theoretical and Computational Nanotechnology*. X:1-33.

Shavelson, R. J., Phillips, D. C., Towne, L. & Feuer, M.J. 2003. On the science of education design studies. *Educational Researcher,* 32(1):25–28.

Shaw, J. & Del Favero, D. 2004. *Advanced Visualisation and Interaction Environment* [Online]. Available from: http://www.icinema.unsw.edu.au/projects/prj_mining.html [Accessed: 12/02/2010].

Shneiderman, B. 1998. *Designing the User Interface: strategies for Effective Human-Computer Interaction.* 3rd ed. New York: Addison Wesley.

Shneiderman, B. & Plaisant, C. 2005. *Designing the User Interface: Strategies for Effective Human-Computer Interaction.* 4th Ed. New York: Addison-Wesley.

Shuttleworth, M. 2009. *External Validity* [Online]. Available from: https://explorable.com/external-validity [Accessed: 21/05/2014].

Simon, H.A. 1981. *The Sciences of the Artificial*. Second Edition. Cambridge, MA. MIT Press.

Slater, M. & Usoh, M. 1994. Body centered interaction in immersive virtual environments. In: N.M. ThalmannN.M. & D. Thalmann (eds). *Artificial life and virtual reality*, London: Wiley, 125-148.

Slater, M. & Wilbur, S. 1997. A Framework for Immersive Virtual Environments (FIVE): speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, 6:603-616.

Smith, L.B. 2003. Learning to recognize objects. *Psychological Science,* 14:244–250.

Smith-Atakan, S. 2006. *Human-Computer Interaction*. Middlesex University Press:Thomson.

Snyder, C. 2003. *Paper Prototyping: The fast and easy way to design and refine user interfaces*. San Francisco: Morgan Kaufman Publishers.

Soloway, E., Jackson, S.L., Klem, J., Qumtan, C., Reed, J., Spitulnik, J., Stratford, S.J., Studer, S., Eng, J. & Scala, N. 1996. Learning Theory in Practice. Case Studies in Learner-centred Design. *Proceedings of CHI'96,* ACM:189-196.

Soloway, E. & Pryor, A. 1996. The next generation in human-computer interaction. *Communications of the ACM,* 39(4):16.

Soto, V.J. 2013. Which Instructional Design Models are Educators Using to Design Virtual World Instruction? *MERLOT Journal of Online Learning and Teaching*, 9(3), September 2013.

Squelch, A.P. 1996. Human Computer Interaction in Rock Engineering. *SIMRAC Report GEN203*, December 1996.

Squelch, A.P. 1998. Application of Virtual Reality for Hazard Awareness in South African Gold Mines. *PhD Thesis*. University of Nottingham.

Squelch, A.P. 2000. VR as an industrial training and marketing tool. *Conference Proceedings IFIP 2000* [Online]. Available from: www.irv.ufsc.br/IFIP-WG-9.5/proceedings [Accessed: 12/04/2008].

Squelch, A.P. 2001. Virtual reality for mine safety training in South Africa. *SA Institute for Mining and Metallurgy*, 101(4). SAIMM, Johannesburg.

Squelch, A.P. 2005. Current VR Activities. [E-mail to:] Van Wyk, E.A. (vanwykea@tut.ac.za) July 24.

Squires, D. 1999. Usability and Educational Software Design. *Interacting with Computers*, 11(5):463-466.

Squires, D. & Preece, J. 1996. Usability and Learning: Evaluating the Potential of Educational Software. *Computers & Education*, 27(1):15-22.

Squires, D. & Preece, J. 1999. Predicting Quality in Educational Software: Evaluating for Learning, Usability and the Synergy between them. *Interacting with Computers*, 11(5):467-483.

Ssemugabi, S. 2006. *Usability Evaluation of a Web-Based E-Learning Application: a Study of Two Evaluation Methods*. MSc dissertation, UNISA.

Ssemugabi, S. & De Villiers, M.R. 2007. Usability and Learning: A Framework for Evaluation of Web-Based e-Learning Applications. *Proceedings of EdMedia Conference, 2007*.

Ssemugabi, S. & De Villiers, M.R. 2010. Effectiveness of heuristic evaluation in usability evaluation of e-learning applications in higher education. *South African Computer Journal (SACJ)*, 45:26-39.

Stander, D. 2006. Interview with Mine Manager Anglo Platinum Lebowa. (Notes in possession of author).

Stanney, K.M., Mollaghasemia, M., Reevesa, L., Breaux, R. & Graeber, D.A. 2003. Usability Engineering of Virtual Environments(VEs): Identifying Multiple Criteria that Drive Effective VE System Design. *International Journal for Human-Computer Studies,* 58(2003):447–481.

Steuer, J. 1992. Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication*, 42(4):72-93.

Stone, R.J. 2012. Human Factors Guidance for Designers of Interactive 3D and Games-Based Training Systems. 2nd ed. Birmingham, UK: University of Birmingham.

Stone, R.J. & Knight, J.F. 2012. *3D Displays: A Human-Centred Review*. Yeovil, Somerset, UK: University of Birmingham.

Stothard, P.M. 2007. Developing and Deploying Interactive Training Simulations for the Coal Mining Industry. *Proceedings SIMTECT 2007*.

Stothard, P.M., Galvin, J.M. & Fowler, J.C.W. 2004. Development, Demonstration and Implementation of a Virtual Reality Simulation Capability for Coal Mining Operations. *Proceedings of ICCR Conference*. Beijing, China.

Stothard, P.M. & Laurence, D. 2014. Application of a large-screen immersive visualisation system to demonstrate sustainable mining practices principles. *Mining Technology,* 123(4):199-206.

Stothard, P.M., Otto, D., Laurence, D.C., Galvin, J.M. & Zenari, L. 2001. E-minesafe Safety and Training Simulator. *The Integration of Knowledge and Skills to Achieve Safe Human Responses, UMRC Research Report RR10/01 ISBN No. 0 7334 1844 9*.

Stothard, P.M., Squelch, A.P., Van Wyk, E.A., Schofield, D., Fowle, K., Caris, C., Kizil, M., & Schmid, M. 2008. Taxonomy of Interactive Computer-based Visualisation Systems and Content for the Mining Industry - Part 1. *Proceedings of the AUSIMM Future Mining Conference 2008*, Sydney.

Stothard P.M. & Swadling, P. 2010. Assessment of Maturity of Mining Industry Simulation, Technical Note, *AusIMM International Transactions*, 119(2):102-109.

Stothard, P.M. & Van der Hengel, A. 2009. Feasibility study: Development and Demonstration of Virtual Reality Simulation for the BHPB Olympic Dam Site Inductions. *Research report from University of New South Wales to BHP Biliton*. Australia, 19 Nov 2009.

Strothotte, T. & Schlectweg, S. 2002. *Non-Photorealistic Computer Graphics: Modeling, Rendering and Animation*. San Francisco, USA: Morgan-Kaufman.

STS. 2013. *Simulated Training Solutions* [Online]. Available from: www.simts.co.za [Accessed: 22/03/2013].

Stufflebeam, D. L. 2001. Evaluation checklists: Practical tools for guiding and judging evaluations. *American Journal of Evaluation*, 22(1):71-79.

Stufflebeam, D. L. 2011. Meta Evaluation. *Journal of Multidisciplinary Evaluation*, 7(15):99–158.

Su, K., Wang, H., Wu, Y. & Kuo, N. 2013. The Interface Design and Usability Evaluation of Interactive Virtual Reality Navigation System. *Applied Mechanics and Materials,* 302(2013):635-639.

Sundt Connections. 2013. Innovative Use of BIM Helps Mining Project Achieve Safety, Speed and Quality [Online]. Available from: http://www.sundt.com/media/blog/2013/06/05/innovative-use-of-bim-helps-mining-project-achieve-safety-speed-and-quality/ [Accessed: 15/02/2015].

SunriseVR. 2007. *Virtual Chicago* [Online]. Available from: www.sunrisevr.com [Accessed: 01/04/2010].

Sutcliffe, A. & Gault, B. 2004. Heuristic Evaluation of Virtual Reality Applications. Interacting with Computers, 16(2004):831-849.

Sutcliffe, A.G. & Kaur, K.D. 2000. Evaluating the usability of virtual reality user interfaces. *Behaviour and Information Technology*, 19(6):415–426.

Sveistrup, H. 2004. Motor rehabilitation using virtual reality. *Journal of Neuro Engineering and Rehabilitation*, 1(10).

Swadling, P. & Dudley, J. 2004. VRLoader – a Virtual Reality Training System for a Mining Application. [Online]. Available from: www.siaa.asn.au/get/2395362703.pdf [Accessed: 20 April 2010].

Sweller, J. 1999. *Instructional Design in Technical Areas*. Camberwell, Australia: Acer Press.

Sweller, J., Ayres, P. & Kalyuga, S. 2011. *Cognitive Load Theory*. New York: Springer.

Taffinder, N., Sutton, C., Fishwick, R.J., McManus, I.C. & Darzi, A. 1998. Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery: Results from randomized controlled studies using the MIST VR laparoscopic simulator. *Studies in Health Technologies and Informatics,* 50:24-30.

Tata. 2014. *Simulators Training Equipment* [Online]. Available from: http://www.indiamart.com/company/10086892/simulators-training-equipment.html [Accessed: 10/01/2015].

Teixeira, L., Ferreira, C. & Santos, B.S. 2012. User-centered requirements engineering in health information systems: A study in the hemophilia field. *Computer Methods and Programs in Biomedicine*, 106(2012):160–174.

Teräs, H. & Herrington, J. 2014. Neither the Frying Pan nor the Fire: In Search of a Balanced Authentic e-Learning Design through an Educational Design Research Process. *The International Review of Research in Open and Distance Learning*, 15(2), April 2014.

Thornton, M., Marshall, S., McComas, J., Finestone, H., McCormick, A. & Sveistrup, H. 2005. Benefits of activity and virtual reality based balance exercise programmes for adults with traumatic brain injury: Perceptions of participants and their caregivers. *Brain Injuries,* 12:989-1000.

Thoroughtec, 2012. *Cybermine Mining Simulators* [Online]. Available from: www.thoroughtec.com [Accessed: 20/04/2012].

Tichon, J. & Burgess-Limerick, R. 2011. A Review of Virtual Reality as a Medium for Safety Related Training in Mining. *Journal of Health & Safety Research & Practice*, 3(1).

Torano, J., Diego, I., Menendez, M. & Gent, M. 2008. A finite element method (FEM) – Fuzzy logic (Soft Computing) – virtual reality model approach in a coalface longwall mining simulation. *Automation in Construction*, 17(4):413-424.

Träskbäck, M. & Haller, M. 2004. Mixed reality training application for an oil refinery: user requirements. *Proceedings of the 2004 ACM SIGGRAPH International Conference on Virtual Reality Continuum and its Applications in Industry*:324–327.

Tsiatsos, T., Andreas, K., & Pomportsis, A. 2010. Evaluation Framework for Collaborative Educational Virtual Environments. *Educational Technology & Society*, 13(2):65–77.

Unger, R. & Mallet, L. 2007. Virtual Reality in Mine Training. *SME Annual Meeting and Exhibit*, February 25-28, Denver, CO, Society for Mining, Metallurgy and Exploration.

UNSW (University of New South Wales). 2013. *iDome* [Online]. Available from: http://www.icinema.unsw.edu.au/projects/infra_dome.html [Accessed: 12/02/2013].

Uusikylä , P. & Virtanen, P. 2000. Meta-Evaluation as a Tool for Learning. Evaluation, 6(1):50-65.

Vaishnavi, V. & Kuechler, W. 2009. *Design Research in Information Systems, Association for Information Systems* [Online]. Available from: http://desrist.org/design-research-in-information-systems [Accessed 08/04/2011].

Valk, K. 2007. Enhancing Usefulness: Connecting Utility and Usability. *MSc Thesis. Eindhoven, July 2007* [Online]. Available from: http://alexandria.tue.nl/extra1/afstversl/tm/valk2007.pdf [Accessed: 26/02/2013].

Van den Akker, J. 1999. Principles and methods of development research. In: *J. van den Akker, N. Nieveen, R.M. Branch, K.L. Gustafson & T. Plomp (eds). Design methodology and developmental research in education and training*:1-14. The Netherlands: Kluwer Academic Publishers.

Van den Akker, J., Gravemeijer, K., McKenney, S. & Nieveen, N. 2006. Introducing education design research. *Educational design research*:3–7. New York: Routledge.

Van den Akker, J. & Plomp, T. 1993. *Development research in curriculum: Propositions and experiences.* Paper presented at AERA meeting, April 12-16. Atlanta.

Van der Linde, P.L., Wessels, C.H. & Kirakowski, T. 2013. A comparative study of three ICT network programs using usability testing. *SA ePublications*, 12(3):174-187.
Van der Straaten, P. 2002. Interaction Affecting the Sense of Presence in Virtual Reality. *Research Task Final Report, Delft University of Technology* [Online]. Available from: http://graphics.tudelft.nl/~vrphobia/intpres.pdf [Accessed: 12/04/2008].

Van der Sandt, J. 2009. Interview with Training Manager, Marula Platinum Mine. (Notes in possession of author).

Van der Woude, S. 2010. Chamber of Mines. First Watch – Promoting Health and Safety Initiatives in the Resources Sector. *Presentation to Bushveld Complex, 21 March 2010* [Online]. Available from: http://www.ammsa.org.za/admin/events/files/Presentation/FW21%20Sout hAfrica.pdf [Accessed: 02/06/2010].

Van der Woude, S. 2012. *South Africa Mine Safety needs more Improvement* [Online]. Available from: http://www.miningsafety.co.za/newscontent/484/South-Africa-mine-safety-needs-more-improvement [Accessed: 03/12/2014].

Vanhoof, J. & Van Petegem, P. 2010. Evaluating the quality of self-evaluations: The (mis)match between internal and external meta-evaluation. *Studies in Educational Evaluation*, 36(2010):20–26.

Van Krevelen, D.W.F. & Poelman, R. 2010. A Survey of Augmented Reality Technologies, Applications and Limitations. *The International Journal of Virtual Reality*, 9(2):1-20.

Van Merriënboer, J.J.G. & Sweller, J. 2005. Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17(2):147–177.

Van Wyk, E.A. 2006. Improving Mine Safety Training Using Interactive Simulations. *Proceedings of the ED-MEDIA 2006 World Conference on Educational Multimedia, Hypermedia and Telecommunications 2006,* Orlando, Florida, 2454-2459.

Van Wyk, E.A. & De Villiers, M.R. 2008. Usability Context Analysis for Virtual Reality Training in South African Mines. *Proceedings of SAICSIT 2008, Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists*. George.

Van Wyk, E.A. & De Villiers, M.R. 2009. Virtual Reality Training Applications for the Mining Industry. *Proceedings of the ACM Afrigraph Conference 2009*, Pretoria.

Van Wyk, E.A. & De Villiers, M.R. 2014. Applying Design-based Research for Developing Virtual Reality Training in the South African Mining Industry. *Proceedings of SAICSIT 2014, Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists*: 70-81. Pretoria.

Vavoula, G. & Sharples, M. 2009. Meeting the Challenges in Evaluating Mobile Learning: A 3-level Evaluation Framework. *International Journal of Mobile and Blended Learning,* 1(2):54-75.

Venable, J.R. 2010. Design Science Research Post Hevner et al.: Criteria, Standards, Guidelines and Expectations. In: *R. Winter, J.L. Zhao & S. Aier (eds). DESRIST 2010:* 109-123. Springer-Verlag.

Venter, D. P. 2007. *Making use of Simple Animations to Improve on the Quality of Strata Control Training on Harmony Mines* [Online]. Available from: www.ammsa.org.za/ne/files/presentations/strata_training_jun07.pdf [Accessed 25/05/2010].

Vrasidas, C. 2004. Issues of Pedagogy and Design in E-Learning Systems. Proceedings of the 2004 ACM Symposium on Applied Computing. March 2004, Cyprus.

VRSim, 2015. *Education and Research* [Online]. Available from:

http://www.vrsim.net/education_and_research [Accessed: 10/01/2015].

Wang, F. & Hannafin, M.J. 2005. Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development, 53*(4):5-23.

Warlick, D. 2002. *Top 10 Smart Technologies for Schools* [Online]. Available from: http://www.techlearning.com/article/17676 [Accessed: 01/04/2010].

Warsash, 2010. Simulators at Warsash Campus. *Southampton Solent University* [Online]. Available from: http://www.solent.ac.uk/wma/simulators.aspx [Accessed: 07/04/2010].

Webber-Youngman, R.C. 2014. Students need to become 'imagineers', academic urges. *Mining Weekly* [Online]. Available from: http://www.miningweekly.com/article/students-need-to-become-imagineers-academic-urges-2014-09-16 [Accessed: 09/01/2015].

Webber-Youngman, R.C.W. & Van Wyk, E.A. 2009. Improving Safety and Health through Virtual Mining Interactive Simulations. *Proceedings of the SAIMM Hardrock Safe Conference*, September 2009, Sun City.

Webber-Youngman, R.C.W. & Van Wyk, E.A. 2010. Improving Health and Safety through Virtual Training. *Proceedings of 2010 Mine Ventilation Society of South Africa Conference*, May 2010, Kempton Park.

Webber-Youngman, R.C.W. & Van Wyk, E.A. 2011. Accident reconstruction simulations – its potential impact on the prevention of future mine accidents. *Proceedings of The South African Institute of Mining and Metallurgy MineSafe 2011 Conference*, August 2011, Johannesburg.

Webber-Youngman, R.C.W. & Van Wyk, E.A. 2013. Incident reconstruction simulations-potential impact on the prevention of future mine incidents. *Journal of the Southern African Institute of Mining and Metallurgy*, 113(6):519-528.

Wein, W.S., Piccirilli, A., Coffey, M. & Flemming, M. 2000. *Web Course Evaluation Checklist* [Online]. Available from: http:/iitclass.bloomu.edu/webeval/index.htm [Accessed: 14/08/2008].

Weiss, M. 2008. Results-based Interaction Design. *Educause Quarterly, No 4* [Online]. Available from: http://net.educause.edu/ir/library/pdf/EQM0847.pdf [Accessed: 12/01/2012].

Weiss, P.L. & Jessel, A.S. 1998. Virtual Reality Applications to Work. *Work*, 11(3):277-293.

Welch, R.B., Blackman, T.T., Liu, A., Mellers, B.A. & Stark, L.W. 1996. The effects of pictorial realism, delay of visual feedback and observer interactivity on the subjective sense of presence. *Presence: Teleoperators and Virtual Environments*, 5:263-273.

Welman, J.C. & Kruger, S.J. 2001. *Research Methodology*, 2nd ed. Oxford University Press.

Wenhold, M. 2006. Interview with Safety, Health and Environment Manager, Impala Platinum Rustenburg. (Notes in possession of author).

Werning, S. 2009. *Real Wars on Virtual Battlefields: The Convergence of Programmable Media at the Military-Civilian Margin*. Columbia University Press: Transcript Verlag.

Wharton, C., Rieman, J., Lewis, C. & Polson, P. 1994. The Cognitive Walkthrough Method: A Practitioner's Guide. In: *J. Nielsen & R.L. Mack (eds). Usability Inspection Methods*. New York: John Wiley & Sons.

Wilson, J.R. 1997. Virtual Environments and Ergonomics: Needs and Opportunities. *Ergonomics*, 40:1057–1077.

Wilson, J.R. & D'Cruz, M. 2006. Virtual and Interactive Environments for Work of the Future. *International Journal for Human Computer Studies*, 64:158-169.

Wingate, L.A. 2010. Meta-evaluation: Purpose, Prescription, and Practice. *In: International Encyclopedia of Education*. 3rd ed. 765-774.

Winn, W. 1990. Some Implications of Cognitive Theory for Instructional Design. *Instructional Science*, 19(1):53-69.

Winn, W. 1993. *A conceptual basis for educational applications of virtual reality*. (Technical report) University of Washington.

Winter, G. 2002. A Comparative Discussion of the Notion of 'Validity' in Qualitative and Quantitative Research. The Qualitative Report, 4(3 & 4), March 2000 [Online]. Available from: http://www.nova.edu/ssss/QR/QR4-3/winter.html [Accessed: 16/02/2014].

Witmer, B.G. & Singer, M.J. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments,* 7(3):225-240.

Wong, A., Marcus, N., Ayres, P., Smith, L., Cooper, G.A. & Paas, F. 2009. Instructional animations can be superior to statics when learning human motor skills. *Computers in Human Behavior*, 25(2):339–347.

Yair, Y., Mintz, R. & Litvak, S. 2001. 3D-Virtual Reality in Science Education: An Implication for Astronomy Teaching. *Journal of Computers in Mathematics and Science Teaching*, 2001:294.

Yin, R.K. 2002. *Case Study Research: Design and Methods.* 3rd Edition. Newbury Park: Sage Publications.

Yu, P.P. & Chen, J.P. 2014. Research on 3D Visualization and Data Management Integration Modeling of Digital Mine: A Case Study of Eastern Gejiu Sn-Cu Deposit in Yunnan Province, China. *Advanced Materials Research*, 1073:2036-2041.

Zaibon, S.B. & Shiratuddin, N. 2010. Heuristics Evaluation Strategy for Mobile Game-Based Learning. *6th IEEE International Conference on Wireless, Mobile, and Ubiquitous Techniques in Education (WMUTE).* Kaohsiung, Taiwan: IEEE Computer Society, 127-131.

Zaritsky, R., Kelly, A.E., Flowers, W., Rogers, E. & O'Neill, P. 2003. Clinical design sciences: A view from sister design efforts. [Special issue]. *Educational Researcher*, 32(1):32–34.

Zeltzer, D. 1992. Autonomy, Interaction, and Presence. *Presence*, 1(1):127-132.

Zhang, H., Wang, Y., Zhao B., Li, G. & Lou, Y. 2008. Advances in Blended Learning *Lecture Notes in Computer Science*, 5328(2008):155-164.

Zhang, S., Stothard, P.M. & Kehoe, J. 2010. Evaluation of Underground Virtual Environment Training: Is the Mining Simulation or Conventional Power Point more Effective? *SIMTECT.* Brisbane, Australia.

Zhao, L. & Deek, F.D. 2006. Exploratory Inspection – A Learning Model for Improving Open Source Software Usability. *CHI '06 Extended Abstracts on Human Factors in Computing Systems,* 1589-1594. Montreal: ACM Press.

Ziegeler, S., Moorhead, R.J., Croft, P.J. & Lu, D. 2001. The MetVR case study: meteorological visualization in an immersive virtual environment. *Proceedings of the IEEE Conference on Visualization '01*, October 2001. San Diego, CA:489–492.

Ziv, A., Ben-David, S. & Ziv, M. 2005. Simulation Based Medical Education: an opportunity to learn from errors. *Medical Teacher*, 27(3):193-199.

APPENDIX A: Ethical clearance documentation

A-1: Ethical consent letter from UNISA

UNISA college of science, engineering and technology

Mr E Van Wyk
School of Computing (Student)
UNISA
Pretoria

18 February 2011

TO WHOM IT MAY CONCERN

Permission to conduct PhD research project **Ref: 008/EvW/2011**

The request for ethical approval for your PhD research project entitled: "Improving Mine Safety Training Using Virtual Reality" refers.

The College of Science, Engineering and Technology's (CSET) Research and Ethics Committee (CREC) has considered the relevant parts of the studies relating to the abovementioned research project and research methodology and is pleased to inform you that ethical clearance is granted for your study as set out in your proposal and application for ethical clearance.

Therefore, involved parties may also consider ethics approval as granted. However, the permission granted must not be misconstrued as constituting an instruction from the CSET Executive or the CSET CREC that sampled interviewees (if applicable) are compelled to take part in the research project. All interviewees retain their individual right to decide whether to participate or not.

We trust that the research will be undertaken in a manner that is respectful of the rights and integrity of those who volunteer to participate, as stipulated in the UNISA Research Ethics policy. The policy can be found at the following URL:
http://cm.unisa.ac.za/contents/departments/res_policies/docs/ResearchEthicsPolicy_apprvCounc_21Sept07.pdf

Yours sincerely

Prof JH Kroeze
Chair: School of Computing Ethics Sub-Committee

University of South Africa
College of Science, Engineering and Technology
Preller Street, Muckleneuk Ridge, City of Tshwane
PO Box 392 UNISA 0003 South Africa
Telephone + 27 12 429 6122 Facsimile + 27 12 429 6848
www.unisa.ac.za/cset

A-2: Mine consent letter

IMPALA PLATINUM LIMITED

(Reg. Number [illegible])
Registered as an External Company in
the Republic of South Africa

IMPALA PLATINUM

Impala Platinum Ltd.
P O Box [illegible]
Rustenburg 0300
Tel: (014) [illegible]
Fax: (014) [illegible]

3 October 2008

To who it may concern

AUTHORISATION TO CONDUCT RESEARCH

Mr Etienne van Wyk from Tshwane University of Technology is hereby authorised to conduct research regarding virtual reality training at the Rustenburg site of Impala Platinum (Pty) Ltd. The research will be conducted at the training facilities at #1, #8 and #10 Shafts. We hope that this research will contribute towards effectively using simulation and visualisation in training at Implats.

Yours sincerely

Pieter Anderson
Senior General Manager: Rustenburg Operations

APPENDIX B: Research Instruments

B-1: *DEVREF* Heuristic Evaluation Instrument

Heuristic evaluation questionnaire for the evaluation of desktop VR training programmes in the South African mining industry

The purpose of this heuristic evaluation study is to validate the present evaluation framework, and to evaluate the two VR mine-safety training systems:

Look, Stop and Fix (LSF) generic hazards system, and

***Interactive Simulated Geological Conditions (ISGC)*.**

of which details are included below.

Responses will be treated in a confidential manner, preserving the anonymity and confidentiality of participants.

Protocol for completing the heuristic evaluation questionnaire

1. It is important that the evaluator spend substantial time exploring the VR training system before commencing with the actual heuristic evaluation. Ideally, the evaluator should assume the role of typical trainee who would use the program.

2. The evaluation framework consists of four categories of heuristics:
   - Category 1 - Instructional design: includes criteria related to pedagogical effectiveness, learning theories and multimedia learning design.
   - Category 2: General usability: interface design and interaction, and criteria that support the goals of usability.
   - Category 3: System design: criteria specific to the design of Virtual Reality systems.
   - Category 4: Context-related heuristics: criteria related to the content and the application domain.

3. In order to enable the evaluator to judge the appropriateness of the program's usability in an informed manner, please take note of the following background information related to the systems to be evaluated:
   - Application domain: Safety training for the mining industry.
   - Systems to be evaluated: ***Look, Stop and Fix (LSF)*** generic hazards system and ***Interactive Simulated Geological Conditions (ISGC)***.
   - Target audience: LSF training is done by all the lower-level mineworkers who work underground. Typical job positions are rock drill operator, winch operator and panel operators. The prior exposure of these trainees to computer technology ranges from very limited to none at all. The *ISGC* training is done by higher skilled employees such as shift bosses, artisans and mine captains.
   - System objectives: The *LSF* system simulates the underground working areas, incorporating potential hazards that mine workers need to identify and indicating

possible actions that might be followed in response to each hazard. Trainees must learn to spot these potentially hazardous conditions, identify the hazards correctly, and indicate which action/s should be taken to address the situation. The *ISGC* system focuses on the geological conditions that may cause rock falls. Trainees have to identify the conditions correctly and specify the associated risks and control measures for each condition.

- Context of use: Both systems are used for refresher training of workers returning from their annual leave. Successful completion of *LSF* is compulsory before workers are allowed to work underground again. Trainees not scoring 80% after two attempts are sent for re-training.
- Program development status: Both systems are currently in use at a large platinum mine. Annual upgrades are developed, and the results of these evaluations can be used to improve future versions of the systems.

4. After spending time becoming familiar with the program, the evaluator should work through the program from the beginning to conduct the actual heuristic evaluation. Since the programs are lengthy with repetitive structures, representative samples of the programs can be reviewed. The researcher will be on hand as a facilitator to guide, advise and support evaluators in using their time efficiently.

5. The evaluator should rate each heuristic on a 5 point scale, with 1 representing strongly disagree and 5 strongly agree. An X should be made to indicate the selected rating.

6. At the end of each category, provision is made for evaluators to record additional problems they may encounter during the evaluation. A comments section is also available.

7. The evaluator should also take note that, apart from this heuristic evaluation, the user satisfaction of these systems will be evaluated as a separate questionnaire.

Thank you very much for participating in this evaluation exercise.

Mr EA Van Wyk, Faculty of ICT, Tshwane University of Technology (PhD student)
Prof MR de Villiers, School of Computing, UNISA (Supervisor of Mr van Wyk's PhD study)

Evaluation of desktop VR training applications

Expert evaluation

Consent form

Please note that the inputs are purely for academic use, and will not be used for consulting or commercial purposes. No evaluator names or company/institution names will be published or disclosed.

I ____

working as _______________________________________

at ____

in the department/division of ____________________ ____________________

state that I have not been put under any pressure to participate in this evaluation exercise as an expert evaluator. I was approached to conduct an evaluation and have agreed to participate. I am aware that participation is voluntary and that I may withdraw at any time without negative consequences.

I realise that the findings of the evaluation will be used for research purposes and that the findings may be published in academic publications.

- My name, position and institution will not be published. Pseudonyms will be used instead of participant real names and a separate file with real identities will be kept for member checking purposes.

- My inputs will be used purely for academic reasons. The collected data will be used for the current PhD study and for research articles with the same research objectives as the current study. The data will be stored for five years at UNISA, after which it will be shredded.

Signed ____________________________ Date ________________________________

Category 1: Instructional design heuristics

Name of training system evaluated:____

	Criteria	Rating				
1.1	**Clear goals, objectives or outcomes:** The training program makes it clear to the learner what is to be accomplished and what will be gained from its use.					
	1.1.1. There are clear goals, objectives or outcomes for the training program.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.1.2. Clear goals, objectives or outcomes are communicated at the beginning of the training program.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.1.3. The outcomes are measurable.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
1.2	**Instructional assessment**					
	1.2.1. The program provides assessment opportunities that are aligned with the objectives or outcomes.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.2.2. The assessment opportunities will serve to enhance trainees' performance.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
1.3	**Feedback to user responses**					
	1.3.1. The training program provides trainees with constructive and supportive feedback on their performance.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.3.2. The feedback is relevant to the training content.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.3.3. The feedback informs the trainee regarding his level of achievement in the training program.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.3.4. The feedback indicates incorrect responses and provides information on the correct responses.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
1.4	**Motivation and creativity**					
	1.4.1. The system supports intrinsic motivation by providing challenges to trainees and encouragement when errors are made.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

	1.4.2. The program captures the trainee's attention early and retains it throughout.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.4.3. This training program increases trainees' confidence by providing them with reasonable opportunities to accomplish the objectives successfully.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.4.4. The program engages trainees by its relevant content.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.4.5. The program engages trainees by its interactivity.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
1.5	**Differences between individual users**					
	1.5.1. The system takes account of linguistic and cultural differences by allowing trainees to select between different languages.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.5.2. The system caters for trainees with different levels of expertise regarding the content.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.5.3. The system caters for trainees with different levels of computer experience.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
1.6	**Reduction of extraneous processing in working memory**					
	1.6.1. The training program effectively uses signalling to highlight essential issues, such as restating important points, using headings for important points, or stressing them in audio mode.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.6.2. Redundancy is avoided by not presenting unnecessary information.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.6.3. Redundancy and overload are avoided by not reiterating the same material in multiple modes (e.g. the program presents information using pictures and spoken words, rather than presenting it in pictures, spoken words, and printed words).	Strongly Disagree				Strongly Agree
		1	2	3	4	5

1.7	**Fostering of germane cognitive load (Germane cognitive load is the load devoted to the processing, construction and automation of schemas)**					
	1.7.1. The training program supports the formation of mental schema by explaining where newly acquired knowledge fits into the bigger picture.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.7.2. The system encourages encoding of the training content into long-term memory by presenting questions after each learning segment.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.7.3. Sufficient scaffolding support is provided (in the form of hints, prompts and feedback) to help trainees achieve training goals.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.7.4. The training program presents narration in a colloquial conversational style.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.7.5. The training program prompts trainees to link concrete example information for each problem category to more abstract information.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
1.8	**Appropriate intrinsic cognitive load**					
	1.8.1. Working through the training program does not cause trainees to split their attention between multiple sources of visual information.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.8.2. The program enhances retention by presenting information in learner-paced segments, rather than as a continuous presentation.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.8.3. The system effectively supports dual channel processing of simultaneous visual and verbal material.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

Problems encountered:

Please use this page to mention any other problems that could not fit in the space provided. Fill in the number of the section in the left column and write the problem(s) in the right column.

Number e.g. 2	**Other problem(s) found**

Please write any additional comments or elaborations you may have in the space below.

Category 2: General Usability Heuristics

Name of training system evaluated:__________________________________

	Criteria	Rating				
2.1	**Functionality**					
	2.1.1 The interface provides the level of functionality the user requires to complete a task.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.1.2. The interface provides adequate back button functionality to return to a previous screen.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.1.3. Icons, labels and symbols are intuitive and meaningful to trainees, bearing in mind the level of trainee context and experience.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
2.2	**User guidance**					
	2.2.1. The interface provides clear indications of what the next required action will be.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.2.2. Help for operating the program is accessible at any time and appropriate.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.2.3. Trainees receive clear instructions on how to use the training program.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.2.4. Guidance to solve problems is given in the form of examples, diagrams, videos or photo's.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
2.3	**Consistency**					
	2.3.1. There is consistency in the sequence of actions taken in similar situations.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.3.2. There is consistency in the use of images, prompts, screens, menus, colours, fonts and layouts.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.3.3. Objects, options, and permissible actions are visible so that users do not have to remember instructions.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.3.4. Different screens that have similar operations, use similar elements for achieving similar tasks.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

2.4	**Error Correction**					
	2.4.1. Error messages are expressed in plain language.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.4.2. Learners are provided with the necessary help to recover from cognitive errors.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.4.3. Error messages indicate precisely what the problem is and give simple, constructive, specific instructions for recovery.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
2.5	**System Status**					
	2.5.1. The training program keeps the trainee informed about what is going on through constructive, appropriate and timely feedback.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.5.2. For every action taken by the trainee, there is a visual or audio response by the training program so that learners can see and understand the results of their actions.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.5.3. The program responds to actions initiated by the user and there are no surprise actions from the system's side.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
2.6	**Aesthetics**					
	2.6.1. The screens are pleasing to look at.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.6.2. The buttons and selections are of an adequately viewable size.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.6.3. The text is of an adequately viewable sufficient viewable size.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.6.4. There is not too much content or information on the screens.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
2.7	**Error Prevention**					
	2.7.1. The training program is designed in such a way that the learner cannot easily make serious errors.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.7.2. When the learner makes an error, the system responds with an error message.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

	2.7.3. Trainees can recognize situations where errors are due to the way they provided input, and not due to incorrect content in their response.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.7.4. The system is robust and reliable throughout.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
2.8	**Interactivity**					
	2.8.1. The training program uses clear and simple terminologies that support trainees in understanding how to interact with the system.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.8.2. The program provides interactions that support trainees in learning the necessary content.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.8.3. Working through the program requires regular trainee interactivity to maintain attention and facilitate comprehension.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

Problems encountered:

Please use this page to mention any other problems that could not fit in the space provided. Fill in the number of the section in the left column and write the problem(s) in the right column.

Number e.g. 2	**Other problem(s) found**

Please write any additional comments or elaborations you may have in the space below.

Category 3: VR System Design Heuristics

Name of training system evaluated:____

	Criteria	Rating				
3.1	**User control**					
	3.1.1. The user is able to interact with, or control, the virtual environment in a natural manner.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.1.2. Responses from the environment to the participant's control actions and movements, are perceived as immediate or close-to-immediate.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.1.3. The system permits easy reversal of actions.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.1.4. Trainees are able to exit the system at any time when they need to do so.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
3.2	**Multimodal System output / feedback**					
	3.2.1. The effect of the trainee's actions on objects in the virtual environment, is immediately visible and conforms to the laws of physics and the trainee's perceptual expectations.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.2.2. The visual representation of the virtual world maps to the trainee's normal perception of that environment.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.2.3. Distortions are not noticeable in visual images.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.2.4. Audio is integrated seamlessly into user task activity.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.2.5. Audio information is meaningful and timely.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
3.3	**Presence**					
	3.3.1. Users feel as if they are part of the virtual environment and not isolated from it.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.3.2. The virtual environment experience is consistent with similar real-world experiences.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

3.4	**Orientation**						
	3.4.1. Users do not find it difficult to maintain knowledge (or 'awareness') of their location while moving through the virtual environment.	Strongly Disagree				Strongly Agree	
		1	2	3	4	5	
	3.4.2. The virtual environment includes appropriate spatial labels and landmarks to assist user orientation.	Strongly Disagree				Strongly Agree	
		1	2	3	4	5	
	3.4.3. It is clear to the user how to exit the virtual environment.	Strongly Disagree				Strongly Agree	
		1	2	3	4	5	
3.5	**Navigation**						
	3.5.1. Is it easy for users to move and reposition themselves in the virtual environment.	Strongly Disagree				Strongly Agree	
		1	2	3	4	5	
	3.5.2. Ways of navigating are consistent throughout the system.	Strongly Disagree				Strongly Agree	
		1	2	3	4	5	
3.6	**Object interaction: selection and manipulation**						
	3.6.1. Input devices are easy to use and easy to control.	Strongly Disagree				Strongly Agree	
		1	2	3	4	5	
	3.6.2. Object interactions are designed realistically to reproduce real-world interaction.	Strongly Disagree				Strongly Agree	
		1	2	3	4	5	
	3.6.3. The system provides the ability to rotate 3D objects and increase detail levels when necessary.	Strongly Disagree				Strongly Agree	
		1	2	3	4	5	
3.7	**Fidelity**						
	3.7.1. The simulations in the system are accurate.	Strongly Disagree				Strongly Agree	
		1	2	3	4	5	
	3.7.2. The objects in the virtual environment move in a natural manner.	Strongly Disagree				Strongly Agree	
		1	2	3	4	5	
	3.7.3. The virtual environment displays adequate levels of realism.	Strongly Disagree				Strongly Agree	
		1	2	3	4	5	
	3.7.4. High-fidelity graphics are used where required.	Strongly Disagree				Strongly Agree	
		1	2	3	4	5	

3.8	Various user modes					
	3.8.1. The system employs various user modes to cater for a range of users from novices to experts.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.8.2. The system provides various user-guidance modes, e.g. Free mode, Presentation mode, Guided mode and Discovery mode.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

Problems encountered:

Please use this page to mention any other problems that could not fit in the space provided. Fill in the number of the section in the left column and write the problem(s) in the right column.

Number e.g. 2	Other problem(s) found

Please write any additional comments or elaborations you may have in the space below.

____

____

____

____

____

____

Category 4: Context-specific Heuristics

Name of training system evaluated:____

	Criteria	Rating				
4.1	**Authentic tasks**					
	4.1.1. The training system supports particular work practices in the context of their natural work environment.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.1.2. The system is customised according to learner-specific needs and the relevance of the curriculum.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.1.3. The program includes tasks applicable to the actual job context of the trainee.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.2	**Appropriate reference materials**					
	4.2.1. The system includes additional reference materials, providing information to trainees on standard operating procedures used in the application domain.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.2.2. The reference materials included in the system are relevant to the problem scenarios.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.2.3. The reference materials are at a level appropriate to the trainees.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.3	**Comprehensive scope**					
	4.3.1. The learning material in the program covers all the vital aspects relating to the topics being addressed.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.3.2. The training also covers possible consequences of trainees not applying the learning material correctly in their work place.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.4	**Adaptive design**					
	4.4.1. The design of the training system is adaptive to changes in site practices.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.4.2. The system refers to the latest current standard operating procedures.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

	4.4.3. The system randomises assessment details such as questions and multiple choice answers when presenting assessment opportunities to trainees.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.5	**Appropriate record keeping**					
	4.5.1. The system maintains student records and assessment results.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.5.2. The system monitors and displays student progress.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.5.3. The system ensures legal compliance in the application domain by capturing detailed individual performance data.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.6	**Trainee preparedness**					
	4.6.1. Trainees are shown how to use the software prior to doing the training program.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.6.2. PC literacy pre-training is available to trainees not comfortable with using computers for training.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.7	**Relevant subject matter**					
	4.7.1. The subject matter matches the goals and objectives of the training program.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.7.2. The subject matter is presented in an appropriate content structure.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.7.3. The information provided in the program is accurate.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.7.4. The system 'speaks the trainee's language' by using terms, phrases, symbols and concepts familiar to the trainee and common to the application domain.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.7.5. The level of language use, in terms of grammar and style, is applicable to the target audience.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.8	**Understandable and meaningful symbolic representation**					
	4.8.1. Symbols, icons and terminology used to represent concepts and objects are used consistently.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

	4.8.2. Symbols, icons and terminology used are intuitive within the context of the task.	Strongly Disagree ... Strongly Agree 1 | 2 | 3 | 4 | 5
	4.8.3. Metaphors used correspond to real world objects or concepts.	Strongly Disagree ... Strongly Agree 1 | 2 | 3 | 4 | 5

Problems encountered:

Please use this page to mention any other problems that could not fit in the space provided. Fill in the number of the section in the left column and write the problem(s) in the right column.

Number e.g. 2	**Other problem(s) found**

Please write any additional comments or elaborations you may have in the space below.

____

____

____

____

Thank you for your participation

B-2: User Satisfaction Questionnaire

User satisfaction questionnaire for evaluation of the Look, Stop and Fix Virtual Reality training system

*The purpose of this questionnaire is to evaluate the **Look, Stop and Fix (LSF)** generic hazards system. Please answer the following questions. All the information you provide in this questionnaire is confidential and will be used for research purposes only.*

Section 1: Biographic details

Name	
Age	
Job title	
Employee number	
Name of Mine you currently work for	
Number of years at this mine	
Number of years in the mining industry	
Home Language	

Answer the following questions by making an X in the correct box:

1.1. What was the highest standard or grade that you completed at school?

< St 3 (< Grade 5)	St 4 – 5 (Grade 6-7)	St 6 - 7 (Grade 8-9)	St 8 - 9 (Grade 10-11)	St 10 (Grade12)

1.2. Have you used any of the following devices before?

	YES	NO
Computer		
Cell phone		
Bank Automatic Teller Machine (ATM)		

1.3. Did you do the PC Literacy pre-training on how to use the computer mouse before you started the Look, Stop and Fix program?

YES	
NO	

Section 2: The following section contains questions on the Look, Stop and Fix (LSF) training system.

2.1. What language did you select to use in the LSF program?

English	Setswana	Xhosa	Sepedi

2.2. How interesting was this training program to you?

Very interesting	Interesting	Average	Not really interesting	Not at all interesting

2.3. How much did you enjoy doing this program on the computer?

Very much	Much	Average	Not really	Not at all

2.4. How easy was this training program to use?

Very easy	Easy	Average	Difficult	Very difficult

2.5. How easy was it to work with the mouse?

Very easy	Easy	Average	Difficult	Very difficult

2.6. How much assistance did you require from the facilitator?

Very much assistance	Much assistance	Average	Not much assistance	No assistance

2.7. How well could you understand the questions in the program?

Very well	Well	Average	Not well	Not at all

2.8. In the Look, Stop and Fix program, how easily did you recognise the objects on the screen?

Very easily	Easily	Average	With difficulty	With great difficulty

2.9. How much did you learn by using this program?

Very much	Much	Average	Not much	Nothing at all

2.10. How much are you at ease using computers for training?

Very much	Much	Average	Not much	Not at all

2.11. Do you believe that the accidents you saw in the program can really happen?

None of them	Some of them	Most of them	All of them

2.12. Do you believe that the accidents you saw in the program can really happen to you?

None of them	Some of them	Most of them	All of them

2.13. How realistic were the accidents that you saw in this training program?

Very realistic	Realistic	Average	Not really realistic	Not at all realistic

2.14. Were you given enough time to complete the training program?

Very much time	Much time	Average	Not really enough time	Definitely enough time

2.15. Will this training program help you to be more aware of the hazards in the workplace?

Very much more aware	Much more aware	Average	Not really	Not at all

2.16. How satisfied are you with the feedback that you received from the program while you were doing the training?

Very satisfied	Satisfied	Average	Not really	Not at all

2.17. To what extent are the hazards shown in this program relevant to your job?

Very relevant	Relevant	Average	Not really relevant	Not at all

2.18. How much would you like to do training on the computer like this again?

Very much	Much	Average	Not really	Not at all

2.19. Do you think this type of training on the computer is better than just listening to an instructor in the classroom?

Much better	Better	The same	Not better	Not at all

2.20. Please indicate your preferred method of training.

Classroom lecture	Practical	Video	Computer	Lecture and Computer

Section 3: User comments on the Look, Stop and Fix (LSF) training system.

3.1. What do you think are the best features of the program?

____

____

____

____

3.2. What aspects of the program do you think should be improved?

____

____

____

____

3.3. Is there any other training that you would prefer to do on the computer?

____

___________ ____

3.4. Please describe problems you encountered in using the Look, Stop and Fix program. There is space for you to list more than one problem, if you need to.

____

____

____

____

3.5. Do you have any other comments on the Look, Stop and Fix program?

____

____

____

THANK YOU VERY MUCH FOR YOUR PARTICIPATION

Mr EA Van Wyk, Faculty of ICT, Tshwane University of Technology (PhD student)
Prof MR de Villiers, School of Computing, UNISA (Supervisor of Mr van Wyk's PhD study)

Consent form

I, ____ (First name and surname) state that I have not been put under any pressure to participate in this evaluation exercise, and have willingly participated in it. I am aware that participation is voluntary and that I may withdraw at any time without negative consequences.

I realise that the findings of the evaluation will be used for research purposes and that the findings may be published in academic publications. My privacy will be protected by not printing my name, position or institution in any such publication.

My answers to these questions will be used for academic reasons only. The data will be stored for five years at UNISA, after which it will be shredded.

Signed ____

Date ____

B-3: *DEVREF* Meta-evaluation Questionnaire

META-EVALUATION OF THE *deVRef* FRAMEWORK

This meta-evaluation questionnaire consists of five sections. In Sections 1 to 4 you are requested to assess the criteria in the four categories of the **desktop Virtual Reality evaluation framework (*deVRef*)**, which was used to evaluate the LSF and ISGC prototype systems, and is intended to be transferable to other desktop Virtual Reality training systems. The four categories are Instructional Design, Virtual Reality Design, System Usability, and Context-specific Design.

We request you to indicate whether the ***deVRef*** framework does indeed evaluate the aspects in the target system that it is intended to. Please answer each question and substantiate your answer with a comment in the space provided. There is also space at the end of each section for further comments, and for specifying possible additional criteria that you believe should be added to the framework. For the sake of clarity, each concept to be evaluated is briefly defined.

Section 5 considers the suitability of heuristic evaluation for desktop VR training systems. In this section, please indicate the extent to which you agree with the various statements.

1. Instructional Design

a. **Feedback**: Users should receive information on what action has been done and what has been accomplished.

Does the framework evaluate **feedback?** Please substantiate your answer.

____

____

b. **Visibility**: The controls for different operations and selections are clearly visible.

Does the framework evaluate **visibility?** Please substantiate your answer.

____

____

c. **Constraints**: This concept refers to ways of restricting the kinds of user interaction that can take place at a given moment.

Does the framework evaluate **constraints?** Please substantiate your answer.

____

____

d. **Consistency**: Interfaces should be designed to have similar operations and use similar elements for achieving similar tasks.

Does the framework evaluate **consistency?** Please substantiate your answer.

____

____

e. **Affordance:** Affordance is the aspect in the design of an object that suggests how the object should be used. Objects and systems should be designed with attributes that support users in knowing how to use them, e.g. scroll bars afford moving up and down.

Does the framework evaluate **affordance?** Please substantiate your answer.

____

____

f. **Contiguity:** Learning is increased when information is presented simultaneously in a multi-modal manner by, for example, associated narrations (audio) and animations/graphics (visual). Information directed to one channel is integrated with information in another channel to support better understanding of the subject matter.

Does the framework evaluate **contiguity?** Please substantiate your answer.

____

____

g. **Learner Control:** When the pace of presentation is controlled by the learner, rather than by the program, then learning is increased.

Does the framework evaluate **learner control?** Please substantiate your answer.

____

____

h. **Signalling:** This concept refers to the highlighting of essential material, such as re-stating important points, using subheadings for emphasis, or stressing aspects by using audio mode.

Does the framework evaluate **signalling?** Please substantiate your answer.

____

____

i. **Personalisation**: Words are presented in conversational style rather than formal style.
Does the framework evaluate **personalisation?** Please substantiate your answer.

____

____

j. **Coherence**: Extraneous load can be minimised by eliminating redundant and irrelevant elements.
Does the framework evaluate **coherence?** Please substantiate your answer.

____

____

If any of the above criteria are not evaluated by the *deVRef* framework, are you of the opinion that those criteria should be included in the framework?

____

____

____

Do you think any other instructional design criteria should be added to the *deVRef* framework? If so, please indicate which criteria should be added.

____

____

____

Do you have any other comments relating to the *deVRef* framework's evaluation of instructional design principles for desktop VR training systems?

____

____

2. Virtual Reality Design

a. Interaction: Interaction should be natural, efficient, and appropriate for target users, domains, and task goals.
Does the framework evaluate **interaction?** Please substantiate your answer.

____

____

b. Navigation: Navigation relates to the processes that allow users to move into positions from which they can perform required tasks. Navigational techniques should be easy to use and not cognitively cumbersome or obtrusive.

Does the framework evaluate **navigation?** Please substantiate your answer.

____

____

c. Object selection and manipulation: This item refers to the selection of virtual objects within an environment to reposition, reorient, or query them.

Does the framework evaluate **object selection and manipulation?** Please substantiate your answer.

____

____

d. Multimodal system output: A main feature of VR technology is that users can be presented with multiple modes of input and output, such as speech, video, and sound.

Does the framework evaluate **multimodal system output?** Please substantiate your answer.

____

____

e. Visual output: The visual interface should support optimization of human visual sensory capabilities. There should be no slight irregularities in a display, such as distortions or lags in the opening of visual images.

Does the framework evaluate **visual output?** Please substantiate your answer.

____

____

f. Auditory output: Auditory cues may be effectively used to augment or in some instances replace visual cues. Sound should be used to enhance perception and increase user performance.

Does the framework evaluate **auditory output?** Please substantiate your answer.

____

____

g. Haptic output: This concept refers to a system that provides a sense of touch and feel to users of a virtual environment.

Does the framework evaluate **haptic output?** Please substantiate your answer.

____

____

h. Presence: Presence may be described as the subjective perception of experiencing oneself as being within a computer-generated environment rather than being in one's actual physical location.

Does the framework evaluate **presence?** Please substantiate your answer.

____

____

i. Engagement: User engagement in a virtual environment should be fostered and sustained, thereby enhancing the sense of presence.

Does the framework evaluate **engagement?** Please substantiate your answer.

____

____

If any of the above criteria are not evaluated by the *deVRef* framework, are you of the opinion that it should be included in the framework?

____

____

____

Do you think any other VR design criteria should be added to the *deVRef* framework? Please indicate which criteria should be added.

____

____

____

Do you have any other comments relating to the *deVRef* framework's evaluation of VR design principles for desktop VR training systems?

____

____

3. **System Usability**

a. **Visibility of the system status**: Users should know where they are within the system.
Does the framework evaluate **visibility of the system status?** Please substantiate your answer.

____

____

b. **Enable frequent users to use shortcuts**: When users become more experienced, they desire to reduce the number of interactions and wish to increase the speed of accomplishing tasks.
Does the framework evaluate **if frequent users are enabled to use shortcuts?** Please substantiate your answer.

____

____

c. **Support internal locus of control**: Users should be in control of the system and the system should respond to actions initiated by the user.
Does the framework evaluate **internal locus of control?** Please substantiate your answer.

____

____

d. **Consistency and standards**: The system should be consistent in that the same words, situations, or actions refer to the same thing.
Does the framework evaluate **consistency and standards?** Please substantiate your answer.

____

____

e. **Error prevention**: Apart from giving good error messages, the system should be designed to prevent errors from occurring.
Does the framework evaluate **error prevention?** Please substantiate your answer.

____

____

f. **Recognition rather than recall**: Objects, actions and options should be visible, so that the user does not need to recall information from one part of the interaction to another.
Does the framework evaluate **recognition rather than recall?** Please substantiate your answer.

____

____

g. **Aesthetic and minimalist design**: System dialogue should contain only the information relevant to the task to be performed by the system.
Does the framework evaluate **aesthetic and minimalist design?** Please substantiate your answer.

____

____

h. **Design dialogues to yield closure:** Sequences of actions should be organised into groups so that the user knows where he/she is at any given time.
Does the framework evaluate **if dialogues are designed to yield closure?** Please substantiate your answer.

____

____

i. **Help users recognise, diagnose, and recover from errors**: The system should give error messages that indicate precisely what the problem is and suggest constructive solutions.
Does the framework evaluate if a system **helps users recognise, diagnose, and recover from errors?** Please substantiate your answer.

____

____

j. **Help and documentation:** The information provided by the system should be easy to search and access, be focused on the user's task, and should list concrete steps to be carried out by the user of the system.
Does the framework evaluate **help and documentation?** Please substantiate your answer.

____

____

If any of the above criteria are not evaluated by the *deVRef* framework, are you of the opinion that it should be included in the framework?

____

____

____

Do you think any other general system usability criteria should be added to the *deVRef* framework? Please indicate which criteria should be added.

____

____

____

Do you have any other comments relating to the *deVRef* framework's evaluation of general system usability principles for desktop VR training systems?

____

____

4. Context-specific Design

a. Learning in real-world contexts: Training programs should include tasks that are applicable to the actual job context of the trainee.

Does the framework evaluate whether **learning is situated in real-world contexts?** Please substantiate your answer.

____

____

b. Corresponding concepts: The system uses terms, phrases, symbols and concepts familiar to the trainee and commonly in use within the application domain.

Does the framework evaluate whether **corresponding concepts** are used**?** Please substantiate your answer.

____

____

c. Appropriate language: The level of language use, in terms of grammar and style, is applicable to the target audience.

Does the framework evaluate **appropriate language?** Please substantiate your answer.

____

____

d. Appropriate record keeping: The system maintains student records and results of assessment.

Does the framework evaluate **appropriate record keeping?** Please substantiate your answer.

____

____

If any of the above criteria are not evaluated by the *deVRef* framework, are you of the opinion that it should be included in the framework?

____

____

____

Do you think any other context-specific design criteria should be added to the *deVRef* framework? Please indicate which criteria should be added.

____

____

____

Do you have any other comments relating to the *deVRef* framework's evaluation of context-specific design principles for desktop VR training systems?

____

____

META-EVALUATION OF THE FRAMEWORK METHODOLOGY

5. To what extent do you agree with the following statements relating to Heuristic Evaluation (HE) as a method for evaluating desktop Virtual Reality training systems?

a. HE is an effective evaluation method for identifying problems in the interaction design.

Strongly Agree	Agree	Maybe	Disagree	Strongly disagree

b. HE is an effective evaluation method that is relatively inexpensive to perform.

Strongly Agree	Agree	Maybe	Disagree	Strongly disagree

c. HE is an effective evaluation method that is relatively easy to perform.

Strongly Agree	Agree	Maybe	Disagree	Strongly disagree

d. HE can result in major improvements to a particular user interface.

Strongly Agree	Agree	Maybe	Disagree	Strongly disagree

e. During a short session, an expert evaluator can identify several usability problems.

Strongly Agree	Agree	Maybe	Disagree	Strongly disagree

f. A small number of experts can identify a range of usability problems.

Strongly Agree	Agree	Maybe	Disagree	Strongly disagree

g. Experienced evaluators can suggest solutions to usability problems that individual users may not pick up.

Strongly Agree	Agree	Maybe	Disagree	Strongly disagree

h. Expert evaluators may be biased due to their strong subjective views and preferences, and this may lead to biased reports.

Strongly Agree	Agree	Maybe	Disagree	Strongly disagree

i. It may be difficult to find evaluators who are experienced in both the specific domain of the system and HCI research.

Strongly Agree	Agree	Maybe	Disagree	Strongly disagree

j. Expert evaluation may not capture the variety of real users' behaviours. For example, novice users may perform unexpected actions that an evaluator might not think of.

Strongly Agree	Agree	Maybe	Disagree	Strongly disagree

Do you have any other comments relating to the use of heuristic evaluation as a method for evaluating desktop VR training systems?

____

Thank you for your participation, please complete the accompanying consent form.

Meta-evaluation of the *deVRef* framework

Consent form

Please note that the inputs are purely for academic use, and will not be used for consulting or commercial purposes. No evaluator names or company/institution names will be published or disclosed.

I ____

working as _______________________________________

at ____

in the department/division of ____________________ ____________________

state that I have not been put under any pressure to participate in this evaluation exercise as an expert evaluator. I was approached to conduct an evaluation and have agreed to participate. I am aware that participation is voluntary and that I may withdraw at any time without negative consequences.

I realise that the findings of the evaluation will be used for research purposes and that the findings may be published in academic publications.

- My name, position and institution will not be published. Pseudonyms will be used instead of participant real names and a separate file with real identities will be kept for member checking purposes.
- My inputs will be used purely for academic reasons. The collected data will be used for the current PhD study and for research articles with the same research objectives as the current study. The data will be stored for five years at UNISA, after which it will be shredded.

Signed ____________________________ Date ________________________________

B-4: Revised *DEVREF* Heuristic Evaluation Instrument

Revised heuristic evaluation questionnaire for the evaluation of desktop VR training programmes in the South African mining industry

Category 1: Instructional design heuristics

Name of training system evaluated:____

<table>
<tr><th></th><th>Criteria</th><th colspan="5">Rating</th></tr>
<tr><td>1.1</td><td colspan="6">Clear goals, objectives or outcomes: The training program makes it clear to the learner what is to be accomplished and what will be gained from its use.</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">1.1.1. There are clear goals, objectives or outcomes for the training program.</td><td colspan="5">Strongly Disagree … Strongly Agree</td></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">1.1.2. Clear goals, objectives or outcomes are communicated at the beginning of the training program.</td><td colspan="5">Strongly Disagree … Strongly Agree</td></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">1.1.3. The outcomes are measurable.</td><td colspan="5">Strongly Disagree … Strongly Agree</td></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr>
<tr><td>1.2</td><td colspan="6">Instructional assessment</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">1.2.1. The program provides assessment opportunities that are aligned with the objectives or outcomes.</td><td colspan="5">Strongly Disagree … Strongly Agree</td></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">1.2.2. The assessment opportunities will serve to enhance trainees’ performance and knowledge.</td><td colspan="5">Strongly Disagree … Strongly Agree</td></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr>
<tr><td>1.3</td><td colspan="6">Feedback to user responses</td></tr>
<tr><td rowspan="2"></td><td rowspan="2">1.3.1. The training program provides trainees with constructive and supportive feedback on their performance.</td><td colspan="5">Strongly Disagree … Strongly Agree</td></tr>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr>
</table>

	1.3.2. The feedback is relevant to the training content.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.3.3. The feedback informs the trainee regarding his level of achievement in the training program.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.3.4. The feedback indicates incorrect responses and provides information on the correct responses.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
1.4	**Motivation and creativity**					
	1.4.1. The system supports intrinsic motivation by providing challenges to trainees and encouragement when errors are made.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.4.2. The system provides encouragement when errors are made.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.4.3. The program captures the trainee's attention early and retains it throughout.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.4.4. This training program increases trainees' confidence by providing them with reasonable opportunities to accomplish the objectives successfully.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.4.5. The program engages trainees by its relevant content.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.4.6. The program engages trainees by its interactivity.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.4.7. The program has a captivating storyline	Strongly Disagree				Strongly Agree
		1	2	3	4	5
1.5	**Differences between individual users**					
	1.5.1. The system takes account of linguistic and cultural differences by allowing trainees to select between different languages.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

	1.5.2. In terms of content, the system caters for novice and knowledgeable trainees.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.5.3. The system caters for trainees with different levels of computer experience.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
1.6	**Reduction of extraneous processing in working memory**					
	1.6.1. The training program effectively uses signalling to highlight essential issues, such as restating important points, using headings for important points, or stressing them in audio mode.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.6.2. Redundancy is avoided by not presenting unnecessary information.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.6.3. Redundancy and overload are avoided by not reiterating the same material in multiple modes (.e.g. the program presents information using pictures and spoken words, rather than presenting it in pictures, spoken words, and printed words).	Strongly Disagree				Strongly Agree
		1	2	3	4	5
1.7	**Fostering of germane cognitive load (germane cognitive load is the load devoted to the processing, construction and automation of schemas)**					
	1.7.1. The training program supports the formation of mental schema by explaining where newly acquired knowledge fits into the bigger picture.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.7.2. The system encourages encoding of the training content into long-term memory by presenting questions after each learning segment.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	1.7.3. Scaffolding support is provided (in the form of hints, prompts and feedback) to help trainees achieve training goals.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

		Strongly Disagree				Strongly Agree
	1.7.4. The training program presents narration in a colloquial conversational style.	1	2	3	4	5
	1.7.5. The training program prompts trainees to link concrete example information for each problem category to more abstract information.	Strongly Disagree 1	2	3	4	Strongly Agree 5
1.8	**Appropriate intrinsic cognitive load**					
	1.8.1. Working through the training program does not cause trainees to split their attention between multiple sources of visual information.	Strongly Disagree 1	2	3	4	Strongly Agree 5
	1.8.2. The program enhances retention by presenting information in learner-paced segments, rather than as a continuous presentation.	Strongly Disagree 1	2	3	4	Strongly Agree 5
	1.8.3. The system effectively supports dual channel processing of simultaneous visual and verbal material.	Strongly Disagree 1	2	3	4	Strongly Agree 5

Problems encountered:

Please use this page to mention any other problems that could not fit in the space provided. Fill in the number of the section in the left column and write the problem(s) in the right column.

Number e.g. 2	**Other problem(s) found**

Please write any additional comments or elaborations you may have in the space below.

____

____

____

Category 2: General Usability Heuristics

Name of training system evaluated:________________________________

	Criteria	Rating				
2.1	**Functionality**					
	2.1.1 The interface provides the level of functionality the user requires to complete a task.	Strongly Disagree 1	2	3	4	Strongly Agree 5
	2.1.2. The interface provides adequate back button functionality to return to a previous screen.	Strongly Disagree 1	2	3	4	Strongly Agree 5
	2.1.3. Icons, labels and symbols are intuitive and meaningful to trainees, bearing in mind the level of trainee context and experience.	Strongly Disagree 1	2	3	4	Strongly Agree 5
	2.1.4. Objects are designed with attributes that support affordance.	Strongly Disagree 1	2	3	4	Strongly Agree 5
2.2	**User guidance**					
	2.2.1. The interface provides clear indications of what the next required action will be.	Strongly Disagree 1	2	3	4	Strongly Agree 5
	2.2.2. Help for operating the program is accessible at any time.	Strongly Disagree 1	2	3	4	Strongly Agree 5
	2.2.3. Trainees receive clear instructions on how to use the training program.	Strongly Disagree 1	2	3	4	Strongly Agree 5

	2.2.4. Guidance to solve problems is given in the form of examples, diagrams, videos or photo's.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.2.5. Help for operating the program is appropriate.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
2.3	**Consistency**					
	2.3.1. There is consistency in the sequence of actions taken in similar situations.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.3.2. There is consistency in the use of images, prompts, screens, menus, colours, fonts and layouts.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.3.3. Objects, options, and permissible actions are visible so that users do not have to remember instructions.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.3.4. Different screens that have similar operations, use similar elements for achieving similar tasks.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
2.4	**Error Correction**					
	2.4.1. Error messages are expressed in plain language.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.4.2. Learners are provided with the necessary help to recover from cognitive errors.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.4.3. Error messages indicate precisely what the problem is and give simple, constructive, specific instructions for recovery.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
2.5	**System Status**					
	2.5.1. The training program keeps the trainee informed about what is going on through constructive, appropriate and timely feedback.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

	2.5.2. For every action taken by the trainee, there is a visual or audio response by the training program so that learners can see and understand the results of their actions.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.5.3. The program responds to actions initiated by the user and there are no surprise actions from the system's side.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
2.6	**Aesthetics**					
	2.6.1. The screens are pleasing to look at.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.6.2. The buttons and selections are of an adequately viewable size.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.6.3. The text is of an adequately viewable sufficient viewable size.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.6.4. There is not too much content or information on the screens.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
2.7	**Error Prevention**					
	2.7.1. The training program is designed in such a way that the learner cannot easily make serious errors.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.7.2. When the learner makes an error, the system responds with an error message.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.7.3. Trainees can recognise situations where errors are due to the way they provided input, and not due to incorrect content in their response.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

	2.7.4. The system is robust and reliable throughout.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
2.8	**Interactivity**					
	2.8.1. The training program uses clear and simple terminologies that support trainees in understanding how to interact with the system.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.8.2. The program provides interactions that support trainees in learning the necessary content.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	2.8.3. Working through the program requires regular trainee interactivity to maintain attention and facilitate comprehension.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

Problems encountered:

Please use this page to mention any other problems that could not fit in the space provided. Fill in the number of the section in the left column and write the problem(s) in the right column.

Number e.g. 2	**Other problem(s) found**

Please write any additional comments or elaborations you may have in the space below.

____

____

____

____

____

Category 3: VR System Design Heuristics

Name of training system evaluated:____

	Criteria	Rating				
3.1	**User control**					
	3.1.1. The user is able to interact with, or control, the virtual environment in a natural manner.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.1.2. Responses from the environment to the participant's control actions and movements, are perceived as immediate or close-to-immediate.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.1.3. The system permits easy reversal of actions.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.1.4. Trainees are able to exit the system at any time when they need to do so.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
3.2	**Multimodal System output / feedback**					
	3.2.1 The effect of the trainee's actions on objects in the virtual environment, is immediately visible and conforms to the laws of physics and the trainee's perceptual expectations.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

	3.2.2. The visual representation of the virtual world maps to the trainee's normal perception of that environment.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.2.3. Distortions are not noticeable in visual images.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.2.4. Audio is integrated seamlessly into user task activity.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.2.5. Audio information is meaningful and timely.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.2.6. The system provides appropriate haptic output.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
3.3	**Presence**					
	3.3.1. Users feel as if they are part of the virtual environment and not isolated from it.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.3.2. The virtual environment experience is consistent with similar real-world experiences.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
3.4	**Orientation**					
	3.4.1. Users find it easy to maintain knowledge (or 'awareness') of their location while moving through the virtual environment.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.4.2. The virtual environment includes appropriate spatial labels and landmarks to assist user orientation.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.4.3. It is clear to the user how to exit the virtual environment.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

3.5	**Navigation**					
	3.5.1. Is it easy for users to move and reposition themselves in the virtual environment.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.5.2. Ways of navigating are consistent throughout the system.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.5.3. Logical barriers are used in areas where physical barriers are absent, but to which users should not be granted access.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.5.4. Users can relocate in the system using a terrain map.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
3.6	**Object interaction: selection and manipulation**					
	3.6.1 Input devices are easy to use and easy to control.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.6.2 Object interactions are designed realistically to reproduce real-world interaction.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.6.3 The system provides the ability to rotate 3D objects and increase detail levels when necessary.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
3.7	**Fidelity**					
	3.7.1 The simulations in the system are accurate.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.7.2 The objects in the virtual environment move in a natural manner.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.7.3 The virtual environment displays adequate levels of realism.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

	3.7.4 High-fidelity graphics are used where required.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
3.8	**Various user modes**					
	3.8.1. The system provides various user-guidance modes, e.g. Free mode, Presentation mode, Guided mode and Discovery mode.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	3.8.2. The system provides shortcuts to frequent users.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

Problems encountered:

Please use this page to mention any other problems that could not fit in the space provided. Fill in the number of the section in the left column and write the problem(s) in the right column.

Number e.g. 2	Other problem(s) found

Please write any additional comments or elaborations you may have in the space below.

Category 4: Context-specific Heuristics

Name of training system evaluated:____

	Criteria	**Rating**				
4.1	**Authentic tasks**					
	4.1.1 The training system supports particular work practices in the context of their natural work environment.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.1.2 The system is customised according to learner-specific needs and the relevance of the curriculum.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.1.3 The program includes tasks applicable to the actual job context of the trainee.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.2	**Appropriate reference materials**					
	4.2.2 The system includes additional reference materials, providing information to trainees on standard operating procedures used in the application domain.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.2.3 The reference materials included in the system are relevant to the problem scenarios.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.2.4 The reference materials are at a level appropriate to the trainees.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.3	**Comprehensive scope**					
	4.3.1. The learning material in the program covers all the vital aspects relating to the topics being addressed.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.3.2. The training also covers possible consequences of trainees not applying the learning material correctly in their work place.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

4.4	**Adaptive design**					
	4.4.1. The design of the training system is adaptive to changes in site practices.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.4.2. The system refers to the latest current standard operating procedures.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.4.3. The system randomises assessment details such as questions and multiple choice answers when presenting assessment opportunities to trainees.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.5	**Appropriate record keeping**					
	4.5.1 The system maintains student records and assessment results.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.5.2 The system monitors and displays student progress.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.5.3 The system ensures legal compliance in the application domain by capturing detailed individual performance data.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.6	**Trainee preparedness**					
	4.6.1. Trainees are shown how to use the software prior to doing the training program.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.6.2. PC literacy pre-training is available to trainees not comfortable with using computers for training.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.7	**Relevant subject matter**					
	4.7.1 The subject matter matches the goals and objectives of the training program.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.7.2 The subject matter is presented in an appropriate content structure.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

	4.7.3 The information provided in the program is accurate.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.7.4 The system 'speaks the trainee's language' by using terms, phrases, symbols and concepts familiar to the trainee and common to the application domain.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.7.5 The level of language use, in terms of grammar and style, is applicable to the target audience.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
4.8	**Understandable and meaningful symbolic representation**					
	4.8.1 Symbols, icons and terminology used to represent concepts and objects are used consistently throughout the program.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.8.2 Symbols, icons and terminology used are intuitive within the context of the task.	Strongly Disagree				Strongly Agree
		1	2	3	4	5
	4.8.3 Metaphors used correspond to real world objects or concepts.	Strongly Disagree				Strongly Agree
		1	2	3	4	5

Problems encountered:

Please use this page to mention any other problems that could not fit in the space provided. Fill in the number of the section in the left column and write the problem(s) in the right column.

Number e.g. 2	**Other problem(s) found**

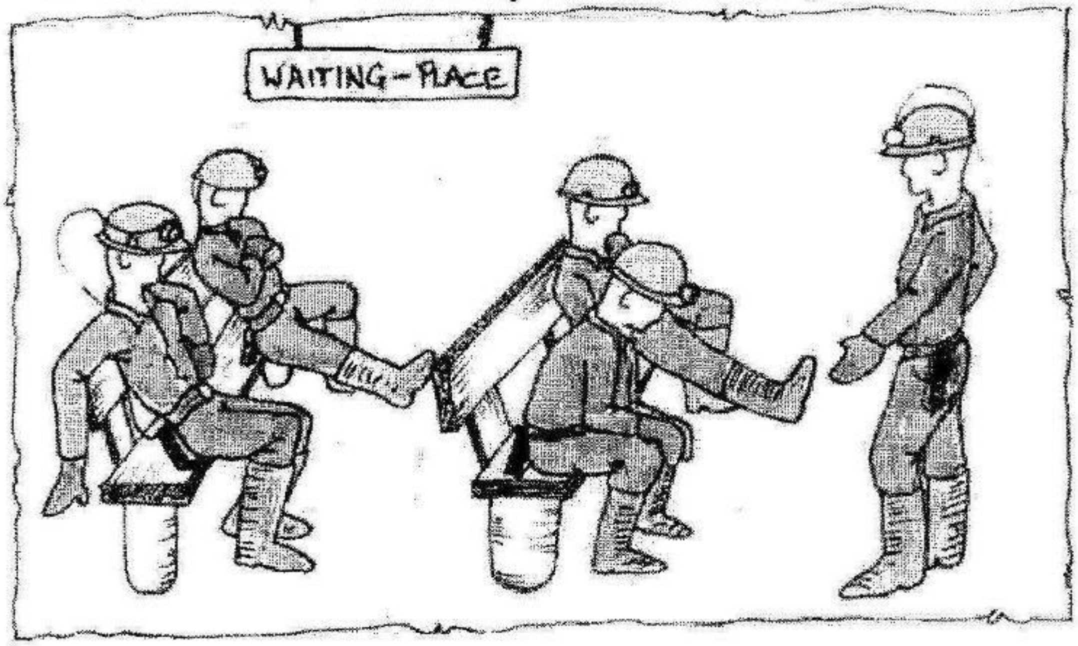
Please write any additional comments or elaborations you may have in the space below.

Thank you for your participation

APPENDIX C: Example Storyboard used in ISGC

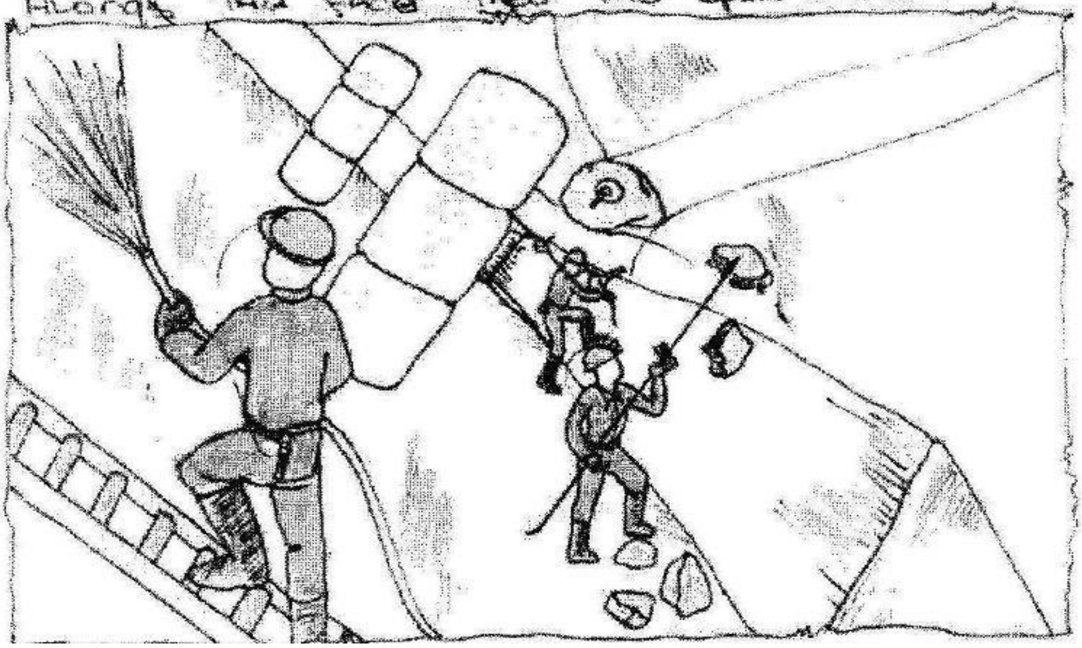
Scene 1.

The Miners and The Team Leader conducting the Daily Safety meeting with the gang



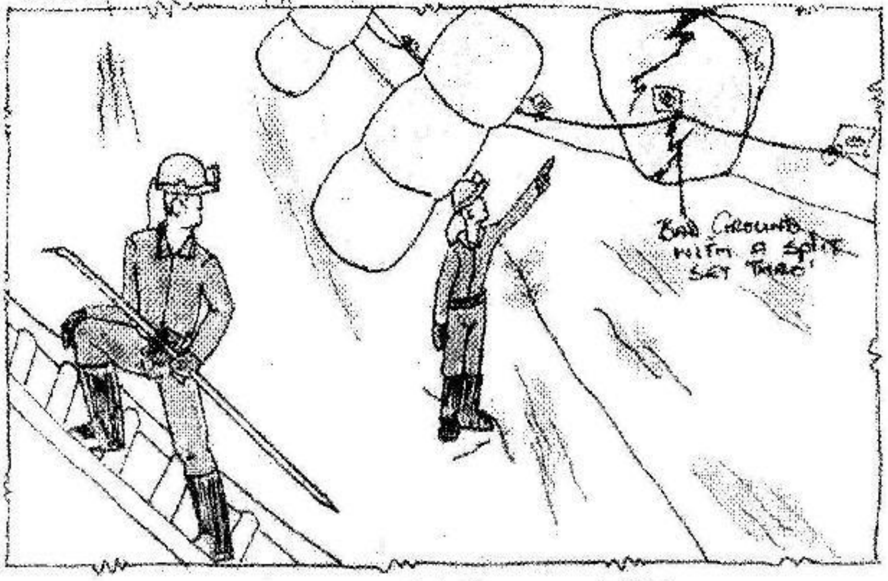
Scene 2

The Team leader and his crew went into The West panel To perform the tasks for The Day. Barring was Done from the top of The panel as well as watering down Safety Lines and slings were being installed Along The face Thro' the split sets.



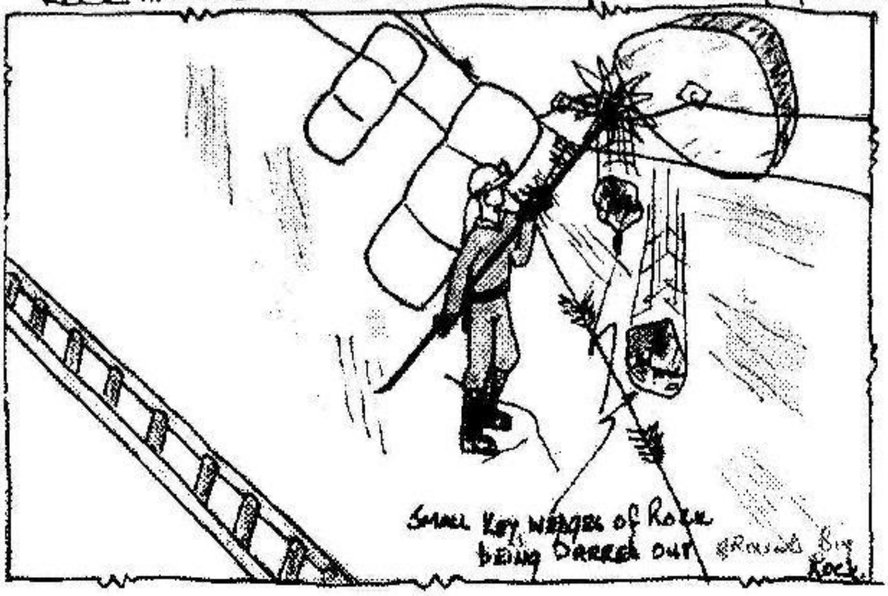
Scene 3

A General Labourer noticed bad ground and pointed it out to the Team Leader, who was busy "making" "safe" above him.



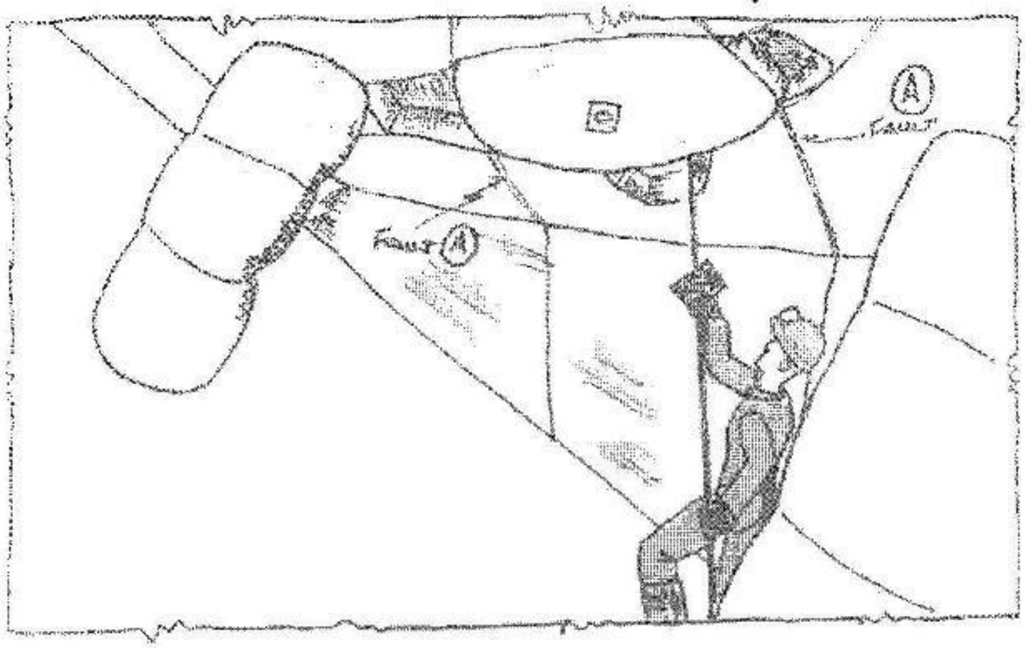
Scene 4.

The Team Leader proceeded to bar the bad Hanging around the Rock-bolted area. Small piece of Rock where barred. The big supported Rock stayed firm. Unaware that he had barred key-wedges from around the large Rock... which also had faults running parallel to the Hanging wall



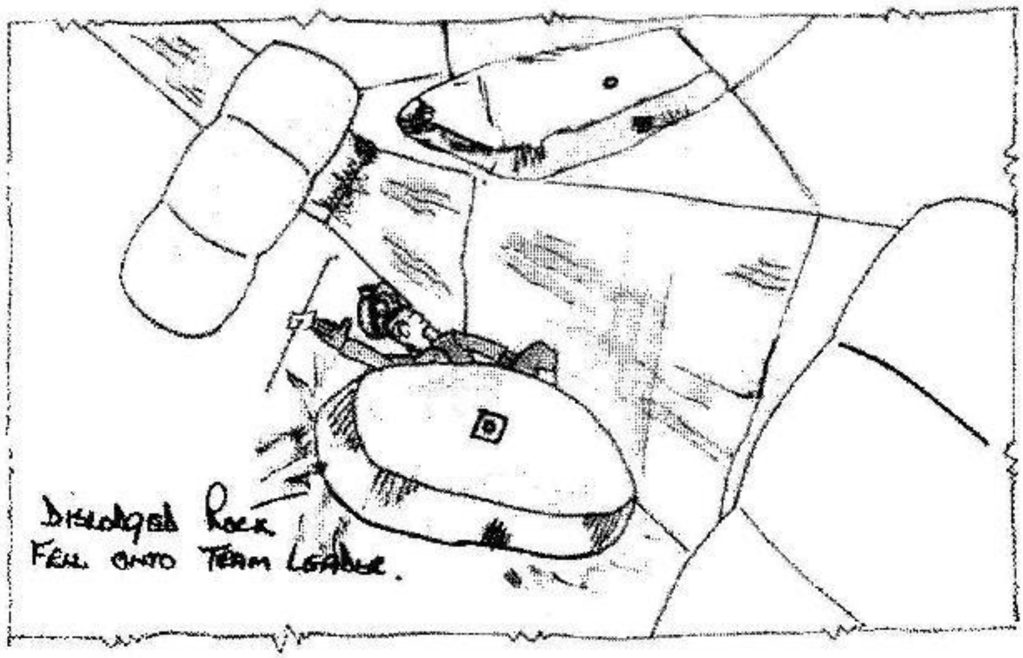
Scene 5.

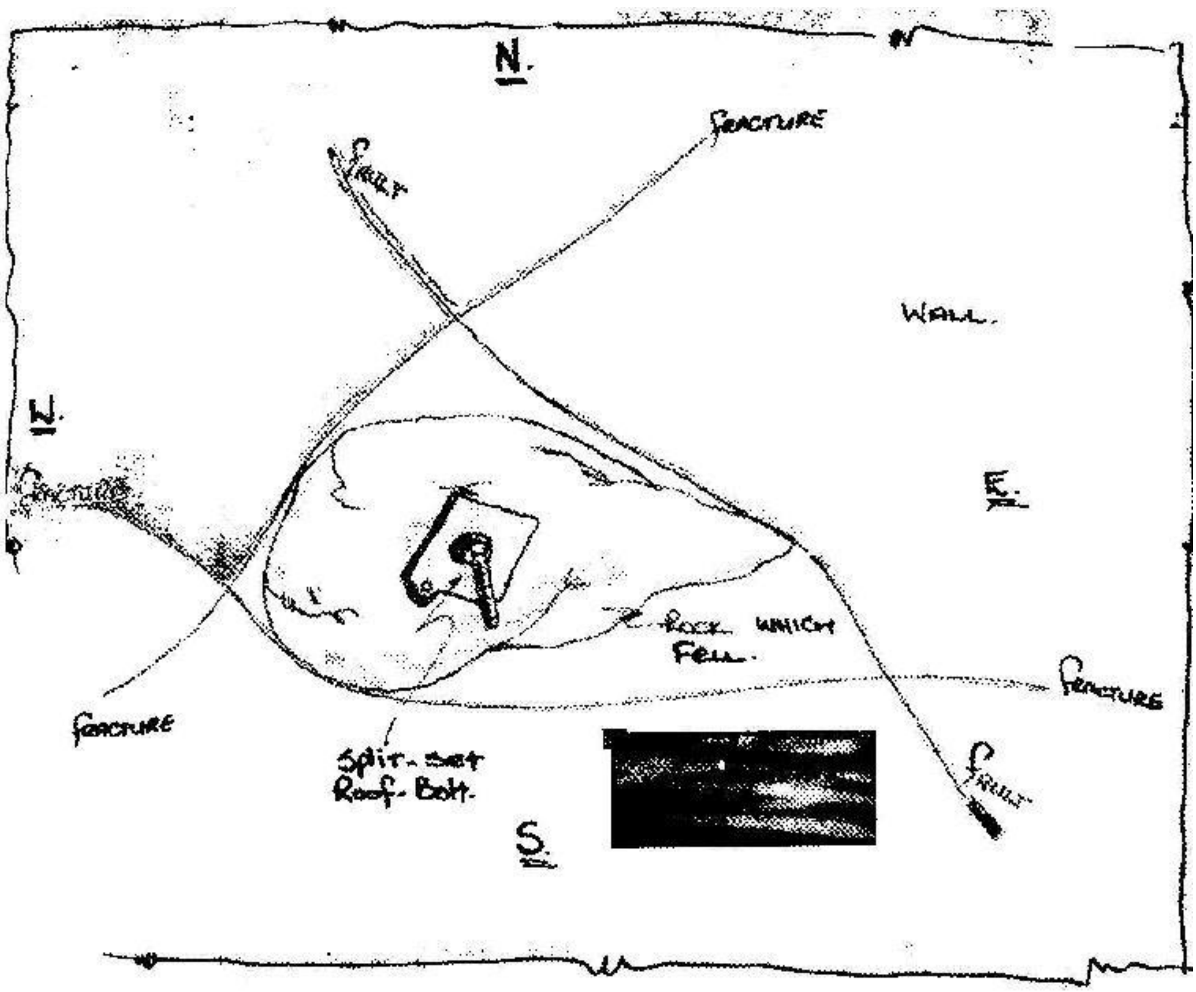
The Team Leader continued Down passed the bad area, barring and sounding as he went.



Scene 6.

When the Team Leader was below the area on the Down-Dip side.... the supported measuring ± 1.5 tons dislodged and struck the Team Leader





NB1 WHEN THE TEAM LEADER SOUNDED THIS ROCK... IT SOUNDED SOLID. HE THEN BARRED AROUND THIS BOULDER, REMOVING KEY WEDGE PIECES, WHICH HOLD THIS ROCK IN PLACE... THUS CAUSING THIS ROCK TO FALL.
THE ROOF-BOLT WAS ALSO TOO SMALL IN LENGTH ONLY ±150mm WAS HOLDING THIS ROCK.

NB2. SHOW THIS TEAM LEADERS SAFETY HARNESS FASTENED TO THIS ROOF-BOLT.

APPENDIX D: Selected publications from this research (included on CD)

D-1: SAICSIT 2008 Conference paper

D-2: AFRIGRAPH 2009 Conference paper

D-3: SAICSIT 2014 Conference paper

D-4: SAIMM 2013 Journal Article