TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF FIGURES

LIST OF TABLES

Preface (Note on writing style)

In this dissertation, I have primarily reported the results from the study in active first person voice to underscore my active involvement and the choices I had to make. I therefore seek the indulgence of the reader in this regard. In addition, I have adopted the following writing style:

Bulleted versus numbered lists

- In lists where I have explicitly stated the number of items in the list, I have used numbered lists. This is to enable ease of referencing to the list at a later stage when required.
- Lists without a specific number of items in the list are provided using bullets.

Writing of numbers as numerals versus words

- Numbers less than ten and all one-word numbers are written in words. For example, the numbers 12 and 30 are written as twelve and thirty respectively.
- All two-word numbers are expressed as numerals. For example, the numbers 71 and 35 are written as is.
- Named entities, such as Chapter 1 and Evaluator 1, are referred to as defined when they are referenced.

CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

This dissertation investigates the applicability of standard usability and accessibility evaluation methods to evaluating a selection of interfaces and educational game applications installed on the Digital Doorway (DD). The study is within the subject of Information Systems. It fits into one of the focus areas of human computer interaction (HCI), namely, usability and accessibility.

In the contemporary information revolution age, the use of information and communication technologies (ICTs) is influencing every aspect of our lives. ICTs are changing the way we teach and learn, the way we conduct business, and how we interact at a social level. However, many people are excluded from the potential economic and social benefits of such new technologies. This phenomenon is generally referred to as the digital divide, that is, the gap between the information technology 'haves' and the 'have-nots' [Bertot, 2003; bridges.org, 2006] (see section 2.3 for a more detailed discussion on the concept of the digital divide).

Several ongoing international and national initiatives aspire to provide access to technologies with the aim of 'bridging' the divide. Examples of such projects include the one laptop per child project [http://www.laptop.org/en/] and the Digital Doorway initiative by the Department of Science and Technology (DST) of South Africa and the Meraka Institute of Council for Science and Industrial Research (CSIR). However, it is sometimes the case that such efforts are directed at the provision of physical ICT devices.

The success of any interactive system, be it an e-commerce site, a company Intranet, or a non-standard[1] computer system like the DD, is dependent on, among other factors, its utility and the ease of use from the user's perspective [Davis, 1989; Nielsen, 2003]. When interactive systems are difficult to use, people will simply stop using them and find other alternatives [Barnum, 2002; Nielsen, 2003].

Usability is generally defined in terms of an application's effectiveness, efficiency, and user satisfaction [International Organization for Standardization, 1998]. Without appropriate usability, content that may be of potential benefit may not be utilized (see section 2.2.1.1 for a more detailed discussion on the concept of usability).

The usability of systems, such as the DD, and the applications installed on them constitute one of the crucial factors to effectively narrow the digital divide [bridges.org, n.d[2]; Nielsen, 2006; Wilson, 2006]. The usability evaluation of interactive systems is one way to determine the potential success

---

[1] Non-standard in this context means systems that do not display standard operating system interfaces or use standard equipments.

[2] Citations with n.d indicate that the references were not dated at the source.

of such systems. It is, therefore, imperative that designers incorporate usability design principles early in the process during the development of projects aimed at narrowing the digital divide. This becomes even more important when the target user group has limited or no computer literacy skills.

1.2 RATIONALE AND MOTIVATION FOR STUDY

The DD project aims to narrow the digital divide by promoting computer literacy through the deployment of DDs in underprivileged communities around South Africa [Gush, De Villiers, Smith and Cambridge, 2010]. The first DD was installed at the rural community of Cwili in the Eastern Cape Province in 2002 and, to date, a total of 206 DDs have been deployed around South Africa.

DDs are aimed at users with little or no computer literacy who may not possess the basic ICT skills to access the software installed on the system. In this type of situation, the system's interface should be particularly supportive and facilitate learning by exploration. It should be tolerant of user error and designers should make every effort to hold the user's attention.

Since its early days, the DD project has primarily focused on providing physical computer access. The systems are deployed without conducting any usability evaluation or applicability tests on the software installed on the system. Given the important role that usability has to play in the effort to narrow the digital divide [bridges.org, n.d; Nielsen, 2006; Wilson, 2006], conducting a usability evaluation of the software installed on the DD is therefore relevant and important.

Physical access is only one aspect of accessibility. The goal of *accessibility* also includes the development of applications that are perceivable, operable and understandable by people with varying abilities, sometimes referred to as disabilities [Henry, 2007] (see section 2.2.1.2 for a more detailed discussion on the concept of accessibility). Although the DD project has focused on hardware development, the hardware does not currently support the use of assistive devices, such as screen readers for visually impaired users. However, the global trend is the enabling of electronic information to people with varying abilities. This is to ensure compliance with regulatory/legal requirements, increase the market share for products and services and for ethical considerations, among others factors [Henry, 2002].

While the DD, in its current form, cannot be evaluated for its accessibility to disabled users requiring access through the use of assistive devices, its level of direct accessibility support can be assessed. Direct accessibility refers to the built-in redundancies in applications which enable as many people as possible to use the application without system modifications [Vanderheiden, 1994]. Direct accessibility support provides benefits not only to users with disabilities, but also to those without.

Furthermore, where the environment of use constrains the use of the system (called situational limitation [Henry, 2002]) direct accessibility support will enhance general usability. It is thus important to consider support for direct accessibility when evaluating the DD.

Given the importance of usability and accessibility in providing effective access to technology, and that there has been no formal usability or accessibility evaluation of the DD software since 2002, this study aims to investigate the applicability of standard usability and accessibility evaluation techniques to evaluating a selection of interfaces and software applications installed on the DD.

1.3 RESEARCH QUESTIONS AND OBJECTIVES

The main research question for this study is:

How can standard usability evaluation methods and accessibility evaluation techniques be used in the evaluation of a non-standard system such as the Digital Doorway?

The main research question is subdivided into three key sub-questions:

1. What evaluation methods are available for evaluating the usability and accessibility of interactive systems?
2. Which of these methods can be applied in the evaluation of the DD?
3. What is the result of applying the suitable method(s) to the DD?

By answering the preceding research sub-questions, my goal is to achieve these three research objectives:

1. To identify evaluation methods applicable to the DD.
2. To develop an instrument that can be used to evaluate the usability and direct-accessibility support provided in the DD.
3. To apply the instrument in the evaluation of the DD in order to assess its effectiveness.

1.4 SCOPE AND LIMITATIONS OF THE STUDY

The study is restricted to usability evaluation of software applications installed on the DD, not the DD hardware. The study evaluates only software applications developed in-house for the DD, so that outcomes will directly influence future development efforts. Thus, third-party applications and programs, such as the Mindset applications, are excluded from the evaluation because the DD team do not have control over the design of these applications. Specifically, the evaluation include: the DD login screen, the registration form for creating a new user account, the main desktop, and three educational games *What-What Mzansi*, *OpenSpell* and *Themba's Journey*. Section 8.2 describes the interfaces and applications evaluated.

In addition, the accessibility evaluation will not consider accessibility support through the use of assistive technologies because the DD does not currently support the use of assistive devices. Hence, evaluation will be confined to the direct accessibility features built into the software applications on the DD and the use of the system *as is*.

Various factors, discussed in section 3.4.1.1, were considered in selecting appropriate evaluation methods for the DD. As a consequence, evaluation uses the heuristic evaluation method primarily, complemented by field usability evaluation through direct observations and questionnaires situated within a natural DD usage environment. Controlled usability testing in a laboratory, for example, is not utilized due to physical and logistical constraints.

1.5 RESEARCH DESIGN AND METHODOLOGY

My research is based on the design research paradigm. It entails two cycles of the design research phases – awareness, suggestion, development, evaluation, and conclusion. The outer cycle encompasses the entire evaluation while the inner cycle involves the development of the heuristics applied in the evaluation of the DD.

The first (outer) design research cycle, involves investigating the literature on methods for evaluating the usability and accessibility of interactive systems to determine the appropriate methods for evaluating the DD. The heuristic evaluation method is used as the primary method of evaluation, with field usability evaluation using direct observations and questionnaires employed as the methods for triangulation. Triangulation in this context is the use of more than one evaluation method and comparison of the results from the methods.

The development phase of the outer design research cycle triggers a second (inner) cycle of design research. Within the inner cycle, existing usability and accessibility principles, guideline and heuristics for the design of interactive systems are examined for their appropriateness in the development of evaluation heuristics for the DD.

Following the development of the heuristics, the effectiveness of the heuristics is tested by five usability and/or accessibility experts who apply them in the heuristic evaluation of a selection of interfaces and applications on the DD. For comparison and triangulation purposes, a field usability evaluation is conducted by direct observation method. Furthermore, participants in the field evaluation are given semi-structured questionnaires to evaluate the DD. The questionnaires are based on a selection of the heuristics developed in this study.

The data from the study, mainly qualitative data, is analyzed using descriptive statistics and text.

The detail of my research design and methodology is provided in Chapter 3.

1.6 ETHICAL CONSIDERATIONS

Ethical considerations are addressed in three ways:

1. A formal ethical clearance was sought and obtained (Appendix A) from the Research and Ethics Committee of College of Science, Engineering and Technology (CSET) at the University of South Africa (UNISA).
2. Study participants are anonymous and the confidentiality of the information they provide is guaranteed. Informed consents were signed by all the expert evaluators (sample consent form for expert evaluators is provided in Appendix B).
3. Formal permission was obtained from the principal of the school where the field evaluation was conducted. The parents/guardians of the participants also signed informed consent forms.

Samples of the letter of permission from the school principal and the informed consent form signed by the parents/guardians of the participants are provided in Appendix C and D.

1.7 THE SIGNIFICANCE AND CONTRIBUTION OF THIS STUDY

As previously stated, the purpose of my study is to investigate the applicability of the standard usability and accessibility evaluation methods to evaluating a selection of interfaces and applications installed on the DD. The DD project aims to narrow the digital divide in South Africa. The usability and accessibility of the applications installed on the system is vital to effectively achieve this objective.

The primary contribution of my study is the set of multi-category heuristics developed for evaluating a non-standard system like the DD. The heuristics are developed through an extensive investigation of existing design principles and guidelines for interactive systems.

A further contribution involves the use of the design research paradigm to develop the heuristics and evaluate the selected interfaces and applications. No other study that has focused on the development of heuristics for a specific domain has used this approach, or has explicitly stated that it has used this approach.

I also extend the eight factors identified by Dix, Finlay, Abowd and Beale [2004] that influence the choice of evaluation methods for interactive systems by identifying the nature of the system for evaluation as an additional factor impacting the decision.

1.8 LAYOUT OF THE CHAPTERS

This dissertation consists of ten chapters. In Chapter 1, I overview the dissertation, describe the rationale and motivation for the study, its scope and limitations, and summarize the research design and methodology. In Chapter 2, I present background to the various concepts that are relevant to the study. This includes an overview of HCI as a field of study; in particular, the concepts of usability, accessibility and their assessment in interactive systems; and a description of the DD. In Chapter 3, I detail my research design. Chapters 4, 5, and 6 constitute the awareness and suggestions phases of the study. In these chapters, I examine the principles and guidelines for the design of usable and accessible interactive systems and also design guidelines for computer-based educational games. I discuss various methods for evaluating the usability and accessibility of interactive systems in Chapters 4 and 5.

In Chapter 7, the development phase of the design research process, I provide the main contribution of this study, namely, the set of multi-category evaluation heuristics for the DD. The results obtained from the heuristic evaluation of the DD through the application of the set of heuristics are presented in Chapter 8. In Chapter 9, I discuss the use of field usability evaluation through direct observations and questionnaires as methods for triangulating the data obtained from the heuristic evaluation. By way of conclusion to the study, in Chapter 10, I summarize the research findings by revisiting the research question and sub-questions. In this chapter I also reflect on the research methodology and my research process, propose the contributions of this study to the body of knowledge and to the DD project, and suggest further research. At the beginning of each chapter, I provide a graphical illustration of the stage of the specific chapter in the dissertation and a chapter map. Figure 1.1 graphically illustrates the layout of the chapters.

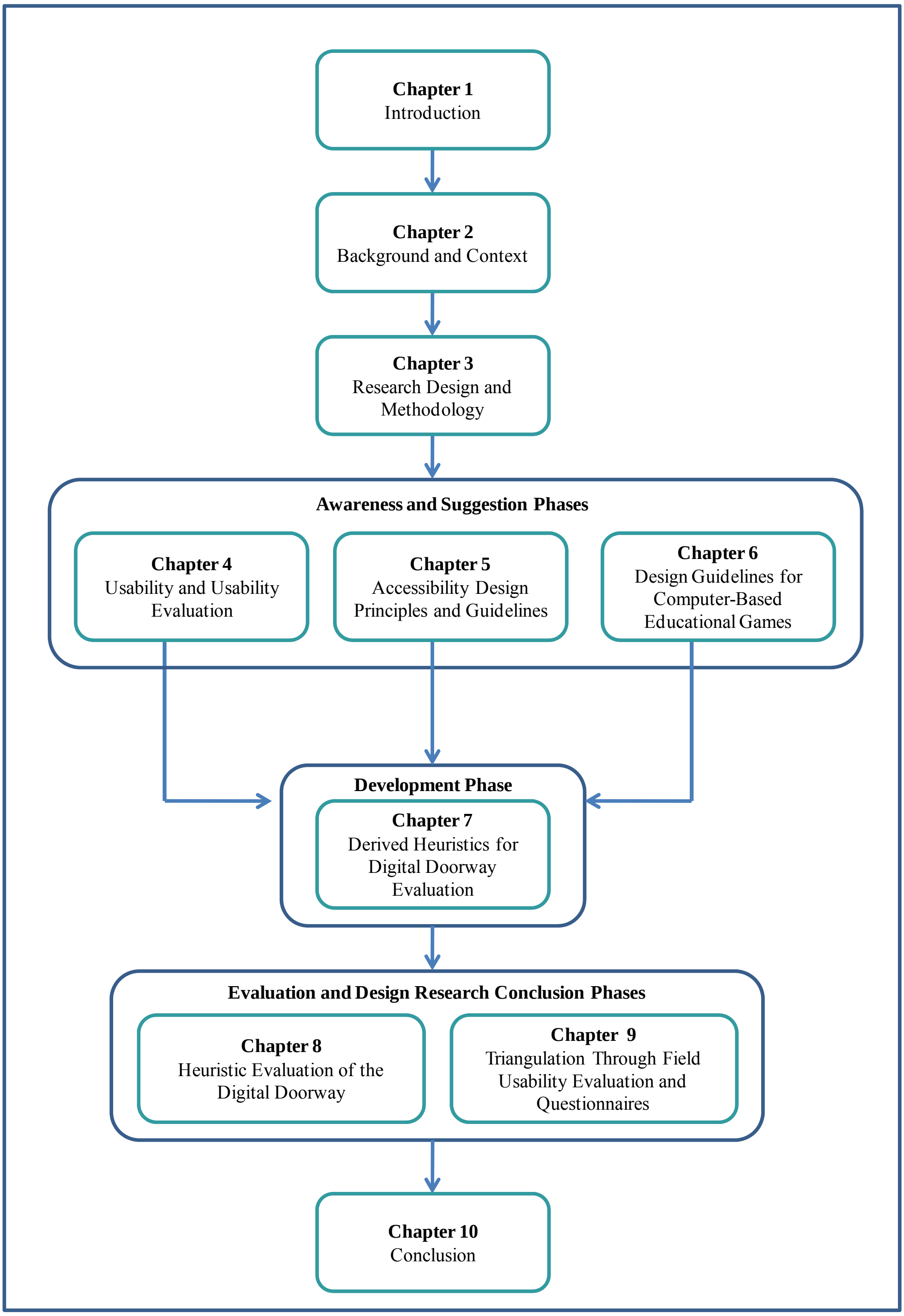
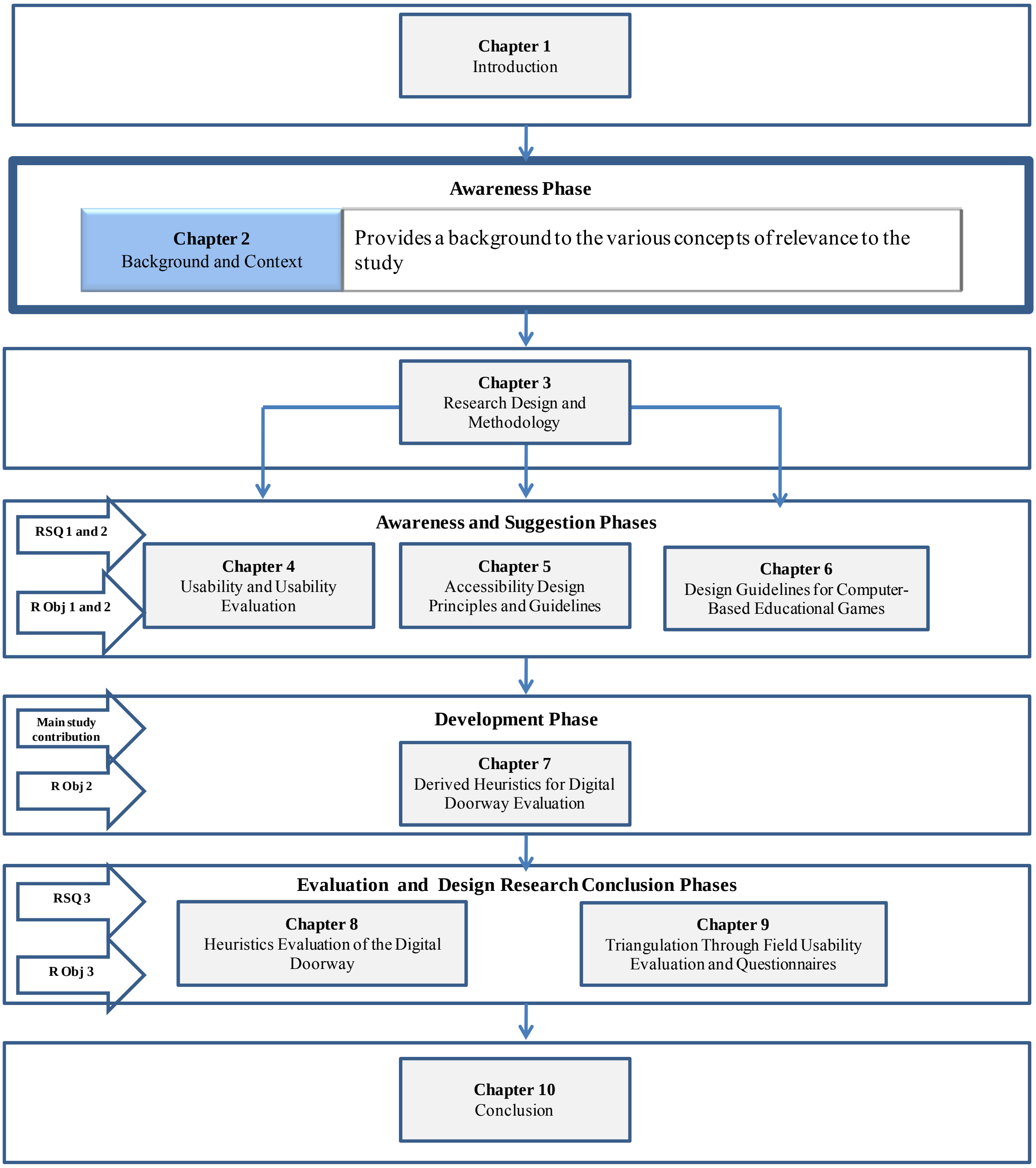


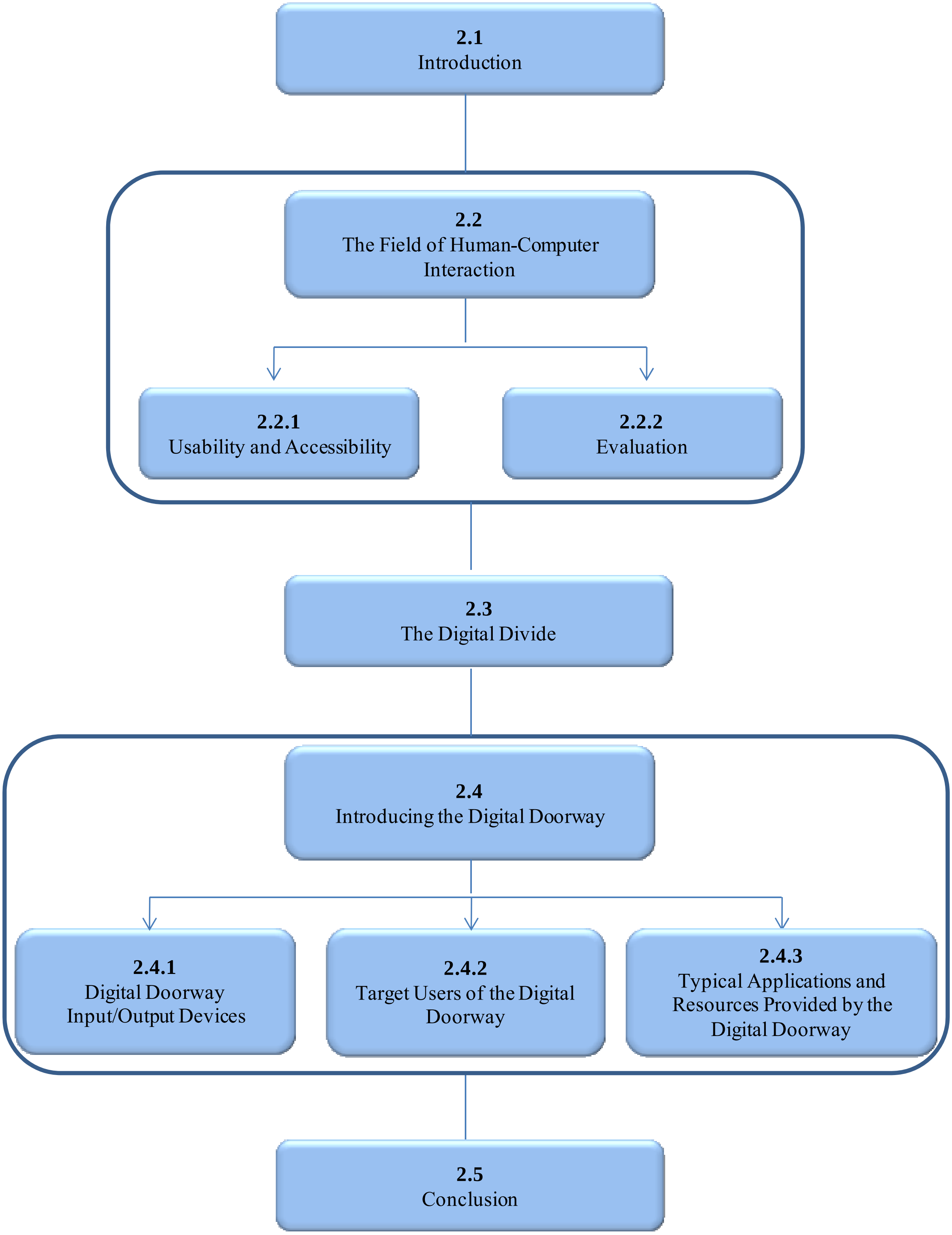
Figure 1.1: Layout of dissertation chapters

CHAPTER 2: BACKGROUND AND CONTEXT

The stage of Chapter 2 in the dissertation



Map of Chapter 2



2.1 INTRODUCTION

In this chapter I provide the background to the various concepts relevant to my study. I introduce the field of HCI, an area of study in which my research falls, in section 2.2. Specifically, I discuss the concepts of usability, accessibility and the evaluation thereof in an interaction system. In section 2.3 I overview the concept of the digital divide together with the various components of the divide. In section 2.4 I introduce, briefly, the Digital Doorway (DD), the system evaluated in this study. I conclude the chapter in section 2.5.

2.2 THE FIELD OF HUMAN-COMPUTER INTERACTION

This study falls within the field of HCI. HCI is a multi-disciplinary field of study concerned with the design, implementation, and evaluation of interactive systems, taking into account the context of use and the task the user needs to accomplish [Dix et al., 2004; Preece, Rogers and Sharp, 2007].

According to Dix et al. [2004], a system *user* can be an individual, a group of people collaborating with one another or a series of users in an organization. The *computer* can be any technology; including desktop computers, large-scale computing systems, process control systems or embedded systems. *Interaction* refers to the communication between the user and the computer, whether direct or indirect. In a direct interaction, there is constant flow of information with feedback and control, for example, editing a document. In an indirect interaction, the user is not directly involved with the interactivity. Examples of this type of interaction include batch processing and low-awareness systems using sensor-based technologies where the lights in a room switch on and remain lit as long as there is movement in the room.

The two main objectives of HCI are [Kotzé and Johnson, 2004]:

1. To enhance the quality of man-machine interaction by systematically applying our knowledge about human capabilities and their limitations as well as the limitations and capabilities of computing devices.
2. To develop or improve productivity and the functionality, safety, effectiveness, and usability of computing systems.

The focus of my study is on usability, one of the domains within HCI. The success of any interactive system is dependent on, amongst other factors, its utility, ease of use from the user's perspective and user experience [Davis, 1989; Dix et al., 2004; Nielsen, 2003]. It is essential, therefore that designers incorporate usability design principles early on in the design process, and evaluate systems for usability before deployment.

This study will also address accessibility as another important element that can affect the quality of human-machine interaction.

Section 2.2.1 introduces the concepts usability and accessibility in more detail, while section 2.2.2 looks at the evaluation of interactive systems with these concepts in mind.

2.2.1 Usability and Accessibility

2.2.1.1 What is Usability?

Earlier definitions of usability were rather simplistic, sometimes equating usability to the presence or absence of certain features or components, such as icons or menus on the interface [Dillon, 2001]. Although the presence of such features may increase the utility of an application, these features may not necessarily be used. Thus, the usability of any application should not be based on such simplistic views but rather on a more comprehensive perspective. For instance, Shackel [1991:24] defined usability as an application's "capability, in human functional terms, to be used easily and effectively by the specified range of users, given specified training and user support, to fulfil the specified range of tasks, within the specified range of environmental scenarios". This definition sets the tone for the formal definition of usability by the International Organization for Standardization (ISO), which defined usability 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" [International Organization for Standardization, 1998].

Other definitions of usability also focus on the user as opposed to the application. For example:

- "Usability is generally regarded as ensuring that interactive products are easy to learn, effective to use, and enjoyable – from the user's perspective" [Preece et al., 2007:20].
- "Usability means that the *people who use the product* can do so *quickly and easily* to accomplish *their own tasks*" [Dumas & Redish, 1993] (as cited by Barnum [2002:6]).

That is, a general theme in each of these definitions is that the usability of an application is really determined by the user's perception of the quality of that application. The user's perception is based on the application's ease of use, the ease of learning and relearning, its intuitiveness to the user as well as the user's perception of the usefulness of the application [Barnum, 2002].

According to Nielsen [1993], usability can be defined in terms of the following quality components:

- *Learnability:* How easy is it for users to accomplish basic tasks the first time they encounter the system?
- *Efficiency:* Once users have learned the use of the system, how quickly can they perform tasks?
- *Memorability:* When users return to the system after a period of not using it, how easily can they re-establish proficiency?
- *Errors:* How many errors do users make, how severe are these errors and how easily can they recover from the errors?
- *Satisfaction:* How pleasant is it to use the system?

Nielsen's [1993] quality components are attributes that can be measured to determine the usability of an application. I discuss the various approaches for evaluating the usability of interactive systems in detail in section 4.2.3.

2.2.1.2 What is Accessibility?

In the past, accessibility concerned only built environment practitioners, for instance, where legislations required cuts into street curbs to enable easy access by people in wheelchairs. However, with the dramatic growth of the World Wide Web (WWW) and the Internet, there is increasing pressure on designers to make electronic information accessible to people with disabilities [Iwarsson and Ståhl, 2003; Rogoff, 2001]. Accessibility is about removing barriers that prevent people with disabilities from taking part in life activities, including the use of services, products and access to information [Bergman and Johnson, 1995]. Thus, in HCI accessibility is defined as the design of applications that "are perceivable, operable and understandable for people with a wide range of abilities" [Henry, 2007].

Advances in ICTs and growth of the WWW have opened new opportunities for people with disabilities. For example, a visually impaired person can now read an electronic version of a newspaper with the help of a screen reader. However, this potential is not realizable if the accessibility barriers impeding people with disabilities in taking advantage of the opportunities offered by new technologies are not addressed [Henry, 2007].

Vanderheiden [1994] distinguishes between direct and indirect accessibility:

- *Direct accessibility* is provided when applications have built-in redundancies that enable as many people as possible to access the application without system modifications or without the use of special adaptive hardware or software. An example of direct accessibility involves the provision of audio feedback in addition to text-based feedback.
- *Indirect accessibility* provides access to electronic information through add-on assistive technologies such as a screen reader.

In Chapter 5, I discuss the ways in which designing can impact on the accessibility of interactive systems to people with various forms of disabilities, together with the methods that are typically employed to evaluate the accessibility of interactive systems.

2.2.1.3 Usability/Accessibility Synergy

Usability is generally defined in terms of an application's effectiveness, efficiency and the satisfaction of the users while accessibility refers to the ability of people with various degrees of disabilities to perceive, operate and understand interactive systems. Some authors consider accessibility to be a subset of usability [Henry, 2002; Ma and Zaphiris, 2003], others [Hudson, 2004; Pühretmair and Miesenberger, 2005] assert that it is a prerequisite for usability. However, what is incontrovertible is that both design concepts contribute to good design. For a given application, the potential users may well be diverse, including people with disabilities. Incorporating usability and accessibility design principles and guidelines will improve general usability for a wider spectrum of users.

In order for an application to be usable, it must be accessible. For instance, a system that provides information through sound alone cannot be used by a user with hearing impairment. While technical accessibility is a prerequisite for accessibility, it is not sufficient to make an application usable. For example, a design that meets the technical specification for the provision of alternative text for graphic elements on an interface may not be usable if the alternative text is not meaningful in the context of use [Henry, 2002].

2.2.2 Evaluation

2.2.2.1 The Role of Evaluation in the Development of Interactive Systems

Design standards, principles and guidelines can direct the design of interactive systems but mere adherence is not enough. To determine whether users will be able to use the application to accomplish real tasks a design should be assessed [Dix et al., 2004]. Evaluation should start early in the design phase and continue throughout the development lifecycle, since it prevents a cascade of changes needed if a late evaluation indentifies problems. The process should also be iterative, with evaluation results being fed back to the next iteration [Dix et al., 2004; Preece et al., 2007]. According to Dix et al. [2004], the goals of evaluation are:

- To determine the ease of use of the systems' functionalities.
- To assess the user interaction experience.
- To identify any specific problem in the system.

Evaluation is beneficial to the following stakeholders [Kotzé and Johnson, 2004]:

- *Designers:* Evaluation enables designers to assess the adequacy of their designs. The result of evaluation can be used in the marketing of an application to convince clients that the product meets their needs. In addition, evaluation results can guide the decision to select between two design alternatives.
- *Clients:* Evaluation enables clients to make an informed decision regarding the application they need to purchase. Evaluation tests can also be used to set milestones in the development process, with practical implementations signed off once they pass the required stages.
- *Users:* The process of evaluation also provides end-users with an opportunity to make their preferences and opinions known to developers. This also helps them to feel that they are part of the development process.

Further, there is an economic reason for early and ongoing evaluation throughout the development process and before the product goes to market; to prevent a disastrous release of bug-ridden application [Preece et al., 2007].

2.2.2.2 Design Principles and Guidelines as a Tool for Evaluating the Usability and Accessibility of Interactive Systems

Interactive systems should be designed to support the user in achieving his/her aims or objectives. Ideally, the composition of design teams should be multi-disciplinary, with practitioners having cognitive science, ergonomics, sociology, and computer science skills, to name a few. This is to ensure that the team understand human problem-solving and physical capabilities, their working environment, and the wider context of interaction [Dix et al., 2004]. However, in reality this is generally not the case. To address this lack of multiple skills requirements, design principles or rules are available that provides designers with a framework that could be followed so as to increase the usability of the end-product.

Various design rules exist, ranging from the most general to application-specific. According to Dix et al. [2004], design rules can be classified according to their level of authority and generality. The authority of a design rule indicates whether, or not, the rule must be followed. The level of generality guides developers in determining whether the given rule can be widely applied in many design situations or if it is applicable to specific types of applications. Dix et al. [2004] identified the following types of design rules:

- *Design principles:* These are high-level, abstract design rules which are widely applicable in a variety of situations. Design principles impose fewer constraints on designers compared to low-level design rules. For instance, the principle of providing adequate feedback can be implemented in a variety of ways [Dix et al., 2004; Kotzé and Johnson, 2004].
- *Design guidelines:* These are less abstract, low in authority and widely applicable. Companies typically have guidelines that ensure their products have the same look and feel [Kotzé and Johnson, 2004].
- *Design standards:* These are specific design rules which are high in authority, but with limited applicability. Standards are typically set by national or international bodies to ensure compliance with them.

Design principles and guidelines, such as the usability principles proposed by Dix et al. [2004], could form the basis for the formulation of heuristics for evaluating the usability and accessibility of interactive systems. *Heuristics* are usability principles or rules of thumb which can be used to guide a design decision or to evaluate interface elements to determine the extent to which they conform to the heuristics [Dix et al., 2004; Preece et al., 2007].

This study involved an extensive literature investigation of existing principles, guidelines and heuristics, for the design of usable and accessible interactive systems, to theoretically ground developing a set of multi-category evaluation heuristics for the DD. Sections 4.3, 4.4.1, 5.5, and 6.5 examine the principles, guidelines and heuristics in detail.

2.3 THE DIGITAL DIVIDE

While computing technology usage was restricted to researchers and scientists in its early days, the use of computing devices (embedded or otherwise) is now widespread. ICTs and the dissemination of information through electronic devices have permeated every aspect of our lives; changing the way we work and play, the way we conduct business and how we learn. However, in many countries access to technology is still restricted to the privileged few.

The term digital divide is generally used to describe the gap between the information technology 'haves' and the 'have-nots' [Bertot, 2003; bridges.org, 2006]. Wilson [2006] formally defines the digital divide as "the inequality in access, distribution and use of ICTs between two or more populations".

The digital divide is a multidimensional phenomenon, affecting different age and gender groups, communities, races and regions of the world [Camacho, 2005]. The divide exists among different population groups within the same nation. For example, in the United States, White and Asian people are over twenty percent more likely to own computers than their Black and Hispanic counterparts [Cooper and Kugler, 2009].

Closer to home, in South Africa, only two percent of Black households had computers in 2001, compared to 46 percent of White households [Statistics South Africa, 2001][3]. This huge disparity between the Black and White population groups can be attributed to the legacy of apartheid and economic exclusion [Martindale, 2002]. A 2007 community survey, conducted by Statistics South Africa, showed a general increase in the ownership of household computers from 8.6 percent of the population in 2001 to 15.7 percent in 2007 [Statistics South Africa, 2007]. However, this report did not provide a breakdown of household computer ownership among the various population groups.

Although the term digital divide is typically used to refer to lack of physical access to ICT devices, Wilson [2006] maintains that effective access requires a number of factors, including the following:

- *Physical access:* This refers to the formal access to ICT devices such as computers, telephones or the Internet. Increasingly, the Web is becoming an essential portal to access and share information. Lack of access to computers and the Internet is considered to be an important aspect of the divide.
- *Financial access*: This refers to the ability of individuals, communities or governments to acquire ICT devices and sustain payments to service providers. For poor communities, whose primary concern is the ability to feed their families, the acquisition of ICT devices are unaffordable luxuries.
- *Cognitive access*: This is the ability of individuals to determine information need, find the information, process and evaluate the information for its appropriateness, and utilize it in a meaningful way. Such effective interaction requires basic ICT skills.
- *Design access (usability)*: Design access refers to the human-computer interface and encompasses the appropriateness of the hardware and software for the target user groups. Of particular importance is the design for people with special needs such as the disabled to avoid technology exclusion. My study specifically addresses the design access aspect of the digital divide.

[3] The latest official numbers issued by Statistics South Africa.

- *Content access*: This refers to the local and cultural relevance of content. For example, the Internet continues to be dominated by the English language (see Figure 2.1). Indeed, according to the 2010 estimate by the Internet World Statistics [2010], no African language featured in the top ten languages of Internet users. From the perspective of a user in a developing country, content access in the local language is the most critical requirement for bridging the digital divide.

My study involved evaluating the software installed on the DD, a development project that was initiated to address the widening digital divide in South Africa.

2.4 INTRODUCING THE DIGITAL DOORWAY

The DD is a walk-up and use, non-standard computer system deployed amongst underprivileged communities around South Africa as part of the effort to narrow the digital divide. The terminal has a metal keyboard with reinforced touchpad for input [Gush et al., 2010]. The robust housing and metal keyboard help to minimize vandalism. Pre-loaded software applications and content run on the Ubuntu Linux operating system.

The DD project is modelled after the 'hole in the wall' concept pioneered by researchers in India who

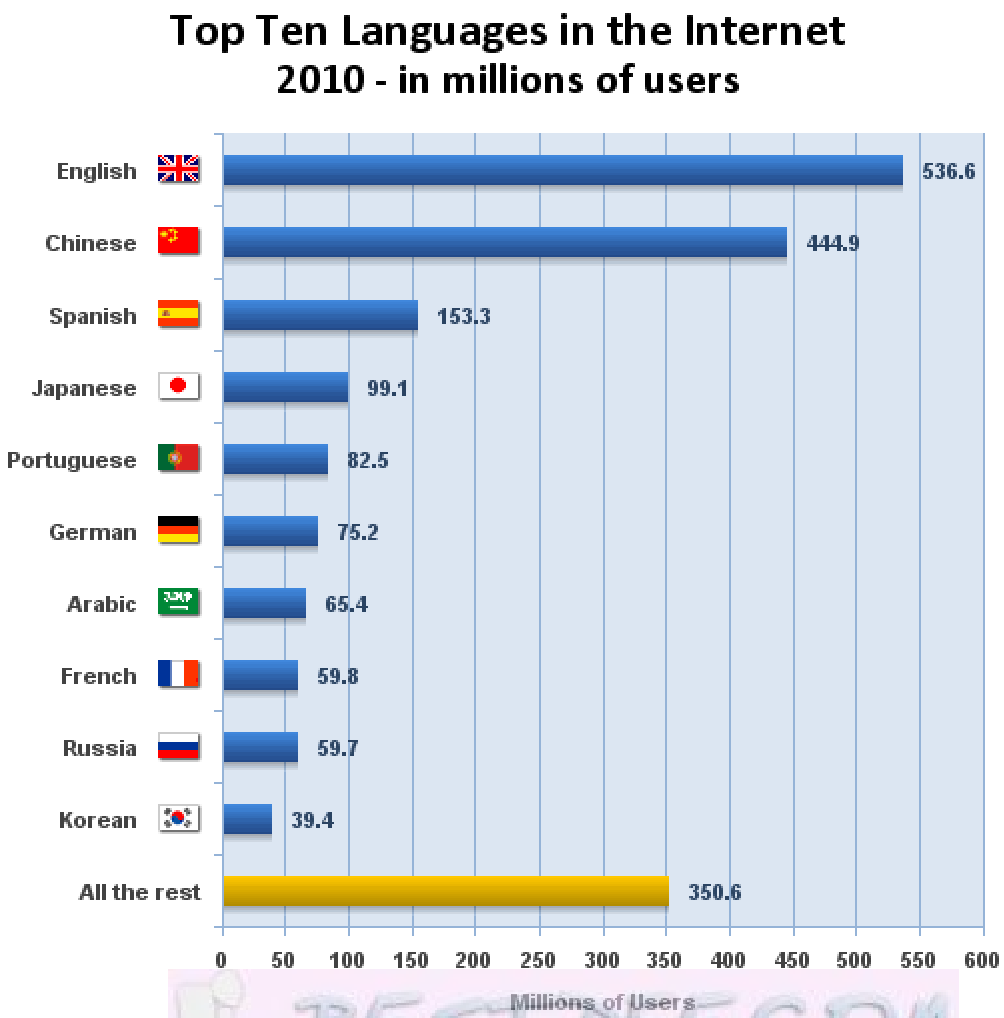


Figure 2.1: Top ten Internet languages [Internet World Stats, 2010]

demonstrated that people, especially children, have the innate ability to acquire basic computing skills through experimentation and unassisted learning. However, independent discovery relies on the availability of technology in an environment that is conducive to this type of experimentation [Mitra and Rana, 2001]. DDs are installed in publicly accessible locations such as schools, police stations and community centres around South Africa.

One of the objectives of the DD project is to narrow the digital divide by providing exposure to computer technology to people in underprivileged communities around South Africa. The first DD was installed at Cwili in the Eastern Cape Province in 2002, with a total of 206 installations to date. Other objectives of the project include [Gush et al., 2010]:

- Testing the viability of unassisted learning as an alternative mechanism for attaining large-scale computer literacy in South Africa.
- Determine whether potential users in a rural community in South Africa would use a personal computer-based outdoor kiosk without any instruction (unassisted learning).
- Provision of technology for social inclusion.
- Provision of meaningful software and content to underprivileged communities.

The first DDs were designed as single terminal units but a need for more standing space led to units with four-terminals. Current DDs are designed with three terminals (for example, Figure 2.2) so as to take advantage of limited space availability while at the same time enabling access by multiple users (see Figure 2.3). One of the terminals acts as a file server for all the terminals. Each of the terminals is equipped with a metal keyboard with reinforced touchpad for input [Gush et al., 2010].

Figure 2.2: A 3-terminal Digital Doorway [http://www.digitaldoorway.org.za/index_main.php?do=hardware]

Figure 2.3 : Multiple users using the Digital Doowaray
[http://www.digitaldoorway.org.za/index_main.php?do=multi]

From 2010, DDs are deployed in self-contained, solar powered units in areas where there is no power supply or suitable venues. A DD container is shown in Figure 2.4.

Figure 2.4: A container Digital Doorway
[http://www.digitaldoorway.org.za/index_main.php?do=hardware]

2.4.1 Digital Doorway Input/Output Devices

The main input device of the DD is a metal keyboard (Figure 2.5). The non-standard metal keyboard provides alpha-numeric keys but no special function keys. There is also a reinforced metal touchpad, which acts as a pointing device.

Pressing the <Enter> key on the keyboard allows the system to accept the user input in the 'username' textbox on the login screen (a detailed description of the login screen is provided in section 8.2.1). The 'Backspace' key is designated for deleting incorrect user input in the registration form (described in section 8.2.2), while the two keys above the metal touchpad produce a 'mouse click' effect when pressed by the user. A tap on the metal touchpad also produce a 'mouse click' effect.

The main output device of the DD is a robust liquid crystal display (LCD) screen (see Figures 2.2 and 2.3), which is covered with plexi glass to prevent vandalism [Gush and De Villiers, 2010]. DDs are not attached to printers, thus users must stand in front of the DD and read content from the screen. Currently, users cannot save information 'downloaded' from the DD to portable storage devices since there are no 'ports' to which these devices can be attached, thus users must write any content onto paper if they wish to take it with them. Further, users cannot upload content onto the DD.

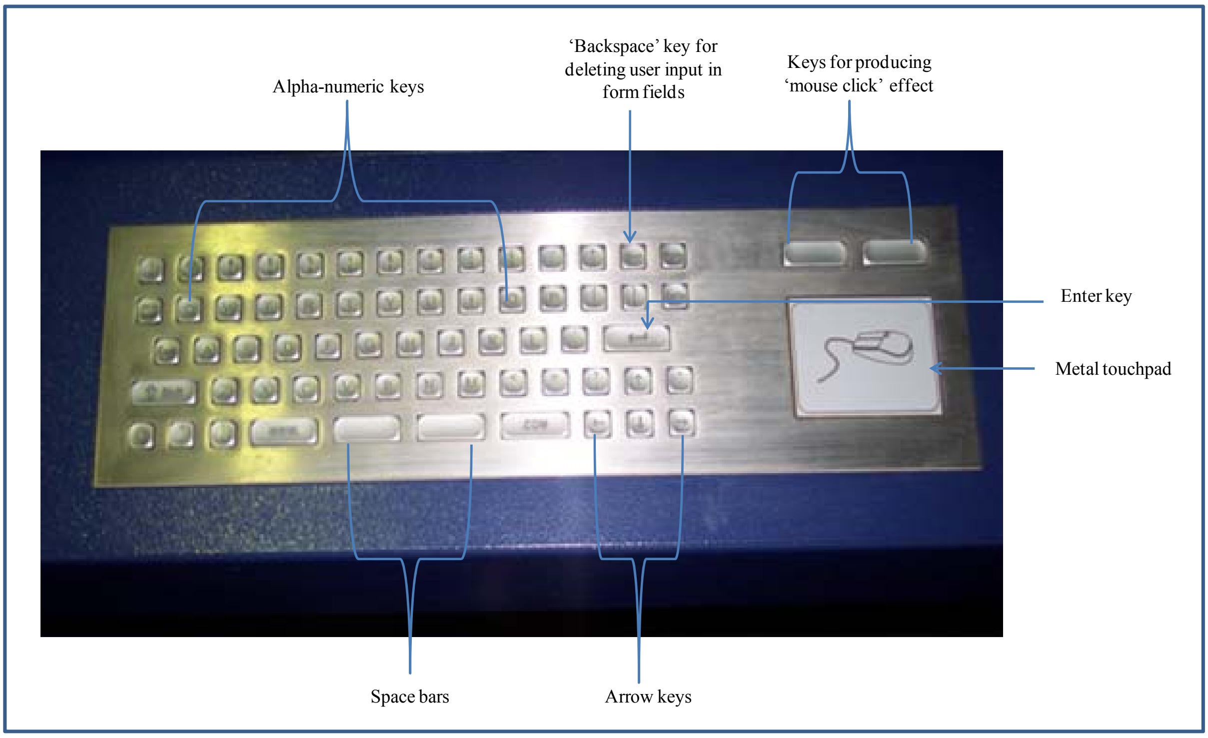


Figure 2.5: Digital Doorway metal keyboard with touchpad

2.4.2 Target Users of the Digital Doorway

DDs are targeted at users with little or no computer literacy in impoverished communities around South Africa. Although the DD project aims to provide access to computing technology to children and adults, the majority of DD users are children and young adults [Greyling and Smith, 2008; Gush et al., 2010]. A recent study on application usage of the DD showed as many as 77 percent of the registered users is male and a mere 23 percent female [Gush and De Villiers, 2010]. DDs are used by single and multiple users, with as many as twelve users sometimes congregating at a single terminal. This allows for individual and peer-assisted learning [Smith, Cambridge and Gush, 2006].

2.4.3 Typical Applications and Resources Provided by the Digital Doorway

The DD provides extensive resources, the majority of these resources are open source or third-party applications [Gush, Cambridge and Smith, 2004]. Resources include educational games, reference materials (Wikipedia), OpenOffice suites, the Mindset applications (a South African curriculum-based educational program), interactive science simulations, and audio books. Applications are preloaded, with content updated on a regular basis [Gush et al., 2010]. Applications that are developed in-house for the DD are typically implemented by contract and visiting developers.

My study evaluates a selection of interfaces and educational game applications developed in-house for the DD system. The specific interfaces and applications evaluated are the login screen, the new user account registration form, the main desktop, and three educational games – *What-What Mzansi* (a quiz game), *OpenSpell* (an educational spelling game) and *Themba's Journey* (for developing life-skills). Detailed descriptions of these interfaces and applications are provided in section 8.2.

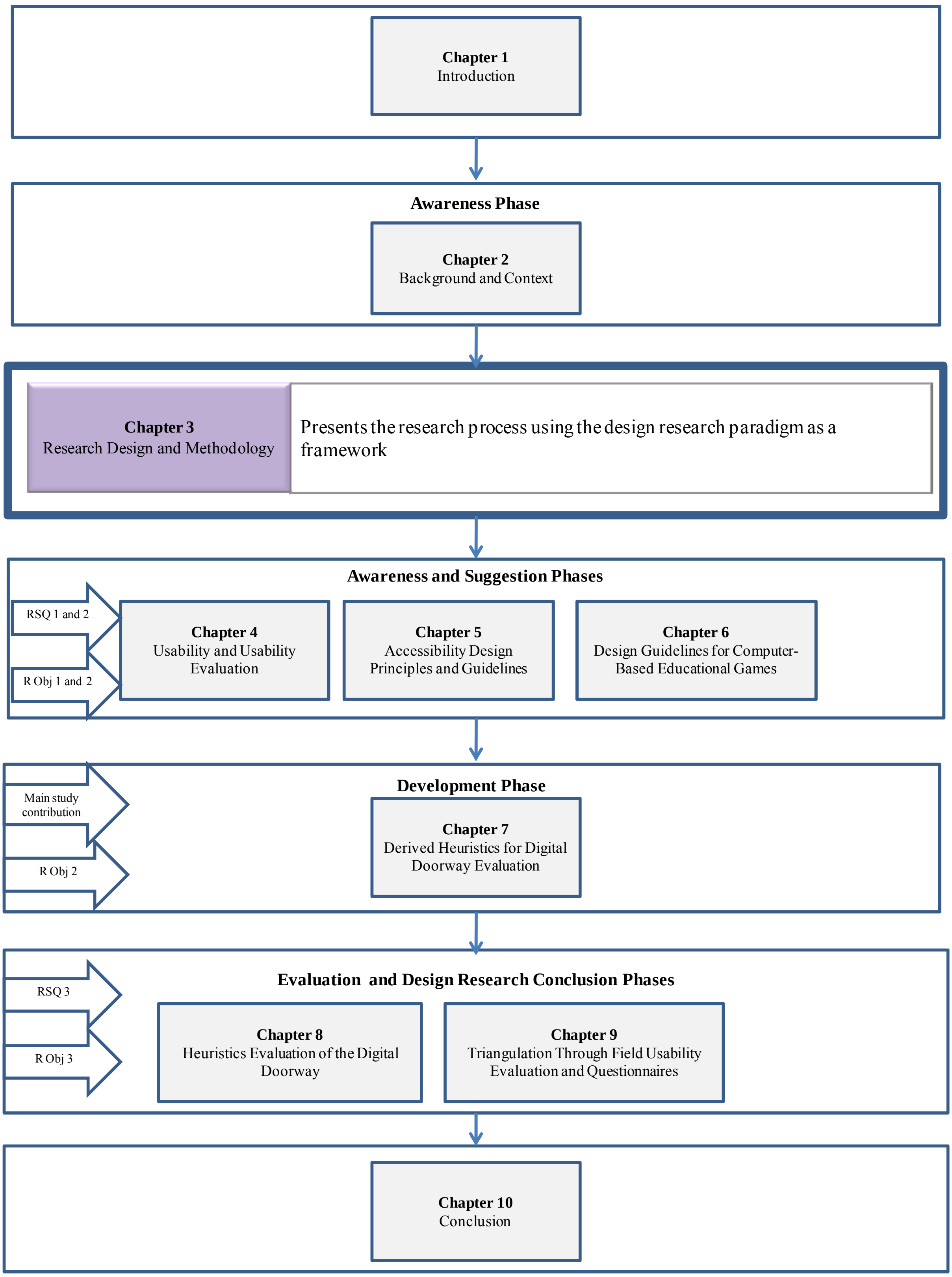
2.5 CONCLUSION

In this chapter I have provided the background to the primary field of study in which my research falls and introduced the specific areas of HCI that my study will address. I described formal definitions of usability and accessibility and introduced design principles as a tool for evaluating the usability of interactive systems. Then I overview the target system to be evaluated, the DD.

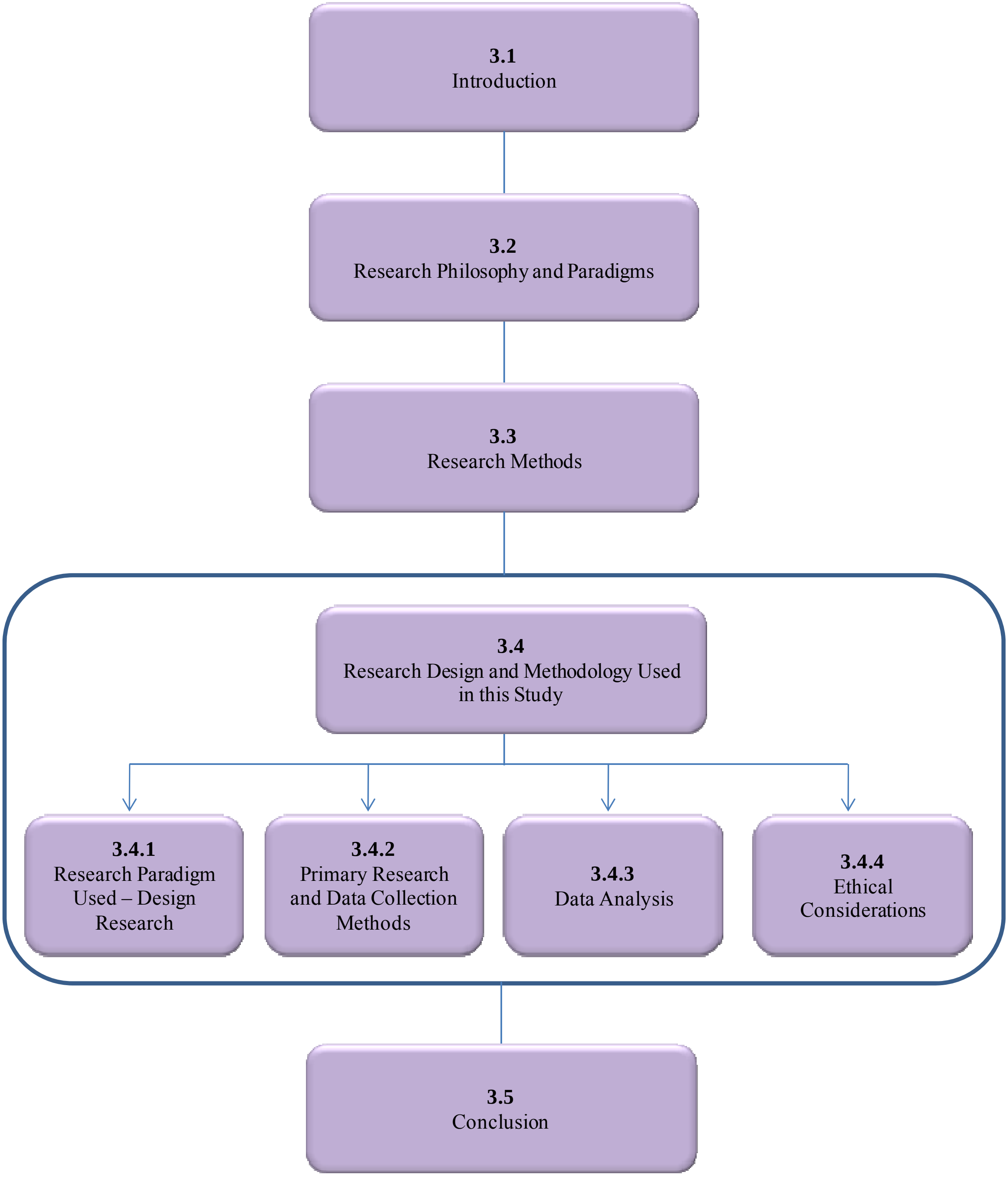
I will explore the various concepts of relevance to my study in the detailed literature investigation of Chapters 4, 5, and 6. However, first I will present the research design and methodology for my study in Chapter 3.

CHAPTER 3: RESEARCH DESIGN AND METHODOLOGY

The stage of Chapter 3 in the dissertation



Map of Chapter 3



3.1 INTRODUCTION

In this chapter I discuss my research design and methodology. Section 3.2 gives the background to different types of research paradigms and their philosophical assumptions, with special emphasis on the design research paradigm (the paradigm relevant to my study). In section 3.3, I discuss briefly quantitative, qualitative and mixed research methods.

While sections 3.2 and 3.3 introduce research paradigms and methods generally, section 3.4 focuses on the specific research design used in my study and detail my research process. The process involved two cycles of the design research phases: the first (outer) cycle encompassed the complete study of evaluating the DD, and the second (inner) cycle developed heuristics used in evaluating the DD. Then, I introduce the three evaluation methods to obtain primary data for this study and a summary of how the data is analyzed. The section also discusses how ethical considerations of relevance to my study are addressed. I conclude the chapter in section 3.5.

3.2 RESEARCH PHILOSOPHY AND PARADIGMS

Research philosophy refers to the explicit or implicit assumptions made by researchers during the research process [Roode, 2009]. Terre Blanche and Durrheim [2006] identifies three philosophical assumptions for research namely, ontological, epistemological and methodological assumptions.

Ontological assumptions relate to the nature of reality to be studied. Epistemological assumptions refer to the kind of the relationship between the researcher and the phenomenon that is being studied while methodological assumptions relate to the ways in which the researcher can go about investigating and obtaining knowledge on the research subject of interest.

Vaishnavi and Kuechler [2004] identify a fourth type of philosophical assumption, axiological assumptions. These refer to things the researcher believes to be of value in relation to the study. All four types of philosophical assumptions influence the research paradigm(s) that will be followed by the researcher.

The three classical types of research paradigms are positivist, interpretive and constructionist research:

1. *Positivist research:* Positivist researchers typically believe in a fixed and stable reality, independent of the researcher. The object of interest is assumed to possess characteristics that can be measured objectively [Roode, 2009]. Positivists generally employ the quantitative research method, which can take the form of experiments or hypothesis testing [Myers, 1997; Terre Blanche and Durrheim, 2006; Vaishnavi and Kuechler, 2004].
2. *Interpretive research:* This research paradigm is based on the assumption that our knowledge of reality is influenced by social constructions, including language, consciousness and shared meanings [Klein and Myers, 1999]. In contrast to the positivists, an interpretive researcher typically interacts with research participants with the aim of better understanding the study

context rather than to make predictions [Roode, 2009; Terre Blanche and Durrheim, 2006; Vaishnavi and Kuechler, 2004].

3. *Constructionist research:* Constructionist research is based on the assumption that reality is socially constructed, and an individual's construct thereof is influenced by societal norms. The social constructionist researcher seeks to understand the contexts in which people live and work to gain a deep meaning of participants' culture and beliefs. The researcher is not detached from the subjects of study and the researcher's interpretation of event is influenced by his/her personal, cultural and historical experiences [Creswell, 2009; Terre Blanche and Durrheim, 2006]. This paradigm is similar to what is described as *critical research* by Myers [1997] and Oates [2006], where the researcher seeks to identify the power relation that exists amongst people, the causes of conflicts, and how the different types of social, cultural and political dominations affects their construction of social reality.

Another research paradigm that is becoming prominent in HCI research, and is relevant to my study, is the *design research* paradigm [Fallman, 2003; Stolterman, 2008; Zimmerman, Forlizzi and Evenson, 2007]:

- Design research is "a problem-solving paradigm which seeks to create innovations that define ideas, practices, technical capabilities, and products through which the analysis, design, implementation, management, and use of information systems can be effectively and efficiently accomplished" [Hevner, March, Park and Ram, 2004].

The main features of positivist, interpretive, constructionist, and design research paradigms, together with their philosophical assumptions are summarized in Table 3.1.

As the main paradigm for my research is that of design research, section 3.2.1 describes design research in detail.

3.2.1 Design Research

Design research analyses the use and performance of designed artefacts, for example a computer system's interface, with the aim of better understanding and improving the artefact [Vaishnavi and Kuechler, 2004]. Design research fits what Simon [1996] (as cited by Vaishnavi and Kuechler [2004]) termed as science of the artificial, which are man-made objects designed to achieve specific objectives.

Design researcher ontologically assume multiple, contextually-situated world states. Their epistemological assumption is based on the concept of 'knowing through making'. This is a process of construction and circumscription, where an artefact is developed and the behaviour of the artefact represent the interactions between different components. The descriptions of these interaction becomes information and, as long as the artefact behaves predictably, the information is assumed to be true [Vaishnavi and Kuechler, 2004]. A design researcher's philosophical perspective can change according to the design research phase s/he is currently engaged in.

Table 3.1: Philosophical assumptions of four research paradigms [Terre Blanche and Durrheim, 2006; Vaishnavi and Kuechler, 2004]

	Philosophical assumptions			
Research paradigms	**Ontology**	**Epistemology**	**Methodology**	**Axiology**
Positivist	- Single, stable reality - Law-like	- Objective - Detached observer	- Experimental - Quantitative - Hypothesis testing	- Truth - Prediction
Interpretive	- Multiple realities - Socially constructed	- Empathetic - Observer subjectivity	- Interactional - Interpretation - Qualitative	- Contextual understanding
Constructionist	- Socially constructed reality - Discourse - Power	- Suspicious - Political - Observer constructing Versions	- Deconstruction - Textual analysis - Discourse analysis	(Not defined by Terre Blanche and Durrheim)
Design	- Multiple, contextually situated realities	- Knowing through making - Context-based construction	- Developmental - Impact analysis of artefact on composite system	- Control - Creation - Understanding

According to Vaishnavi and Kuechler [2004], design research generally has five phases, as illustrated in Figure 3.1:

1. *Awareness of problem:* The design researcher(s) becomes aware of the problem via information from multiple sources such as academia or industry. The end product of this phase is a proposal for new research.
2. *Suggestions:* The design researcher(s) creatively envisions new functionality, based on new or existing elements. This phase results in a tentative design, which may involve developing prototypes.
3. *Development:* The design researcher(s) creates the artefact based on the tentative design. The implementation technique used is influenced by the artefact being constructed.
4. *Evaluation:* The design researcher(s) evaluates the artefact. This may involve qualitative and quantitative methods, which are used to tentatively explain any deviation from expectations. Vaishnavi and Kuechler [2004] consider this process to be an analytic sub-phase within the evaluation phase. The evaluation results, together with lessons learnt from the development phase feed back into the next iteration, which starts at the second phase. These cycles of suggestions, development and evaluation continues until the end product is assessed as ‘good enough’.
5. *Conclusion:* The final phase of the current research effort culminates in the developing of a satisficing artefact, or producing an artefact with behaviour that is ‘good enough’ but not

necessarily optimal. In the conclusion phase, researchers compile reports and group lessons learnt into those that add to the body of knowledge and those that may lead up to further research.

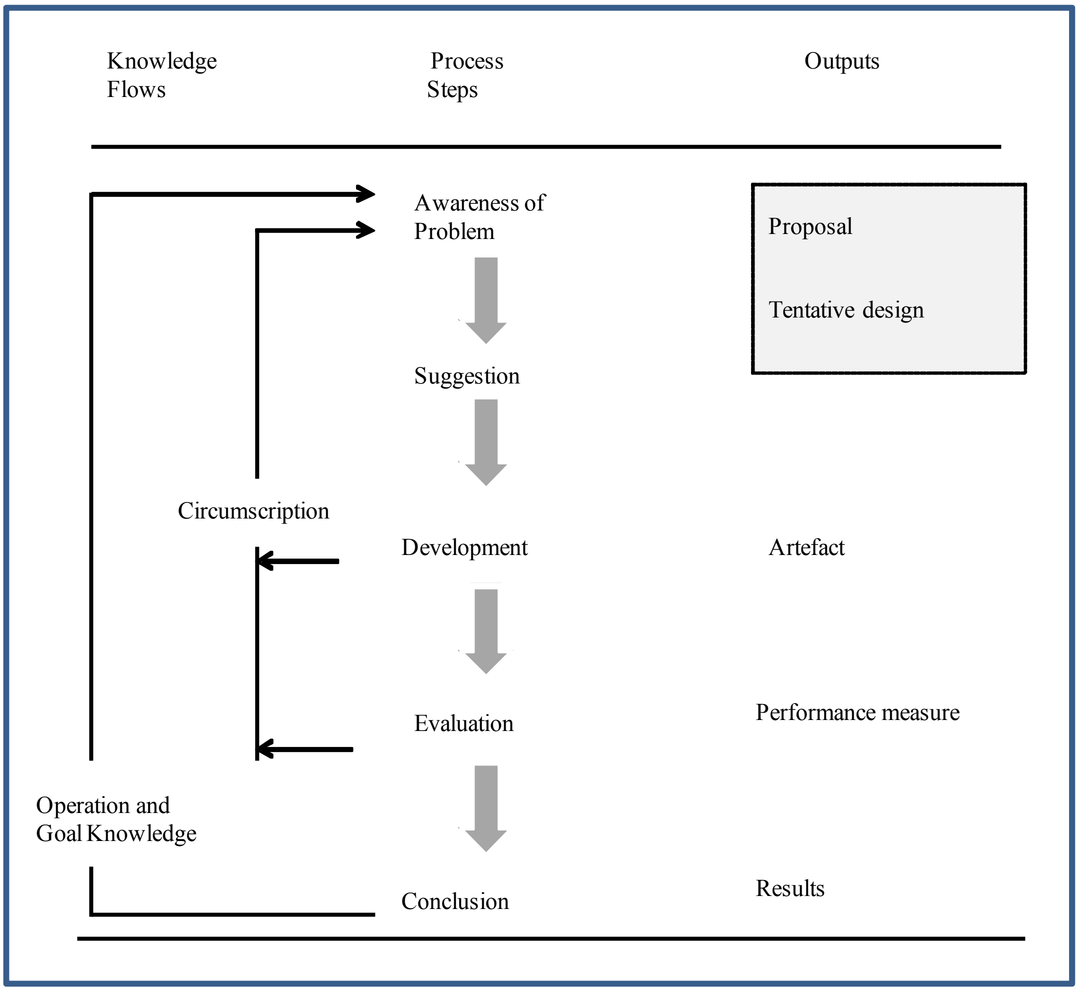


Figure 3.1: Phases of design research

In addition to classifying a research based on the paradigm(s) followed, research methods are commonly classified as quantitative or qualitative, as outline in section 3.3.

3.3 RESEARCH METHODS

Research methods are broadly described as quantitative or qualitative.

- *Quantitative research* methods involve the measurement and analysis of numerical data and the use of statistical packages. Quantitative research methods were originally developed for the study of phenomena in the natural sciences domain [Myers, 1997], but are now widely used in social sciences research. Typical methods to collect data in quantitative studies include experiments, surveys or questionnaires [Myers, 1997; Oates, 2006; Olivier, 2004; Roode, 2009]. Findings from quantitative research can be generalized to the entire population [Oates, 2006; Olivier, 2004].
- *Qualitative research*, which was developed by social sciences researchers, entails using qualitative data to obtain a rich understanding of the phenomenon that is being investigated

[Myers, 1997]. Qualitative studies can be done using methods such as participant observation, interviews, case studies, and action research [Creswell, 2009; Myers, 1997; Oates, 2006; Olivier, 2004]. The data collected in my study are mainly qualitative data.

Rather than being seen as mutually exclusive, quantitative and qualitative research methods can be combined in a single study, termed the *mixed research* method by Creswell [2009]. According to Creswell [2009], the combination of quantitative and qualitative research methods often offset the limitations and biases inherent in the use of any single method and can facilitate a deeper understanding of the area of study. HCI research uses qualitative, quantitative and mixed methods, depending on the nature of the research being done. Further, HCI has evolved certain specialized methods, including heuristic evaluations, usability and accessibility testing, contextual inquiry, user observations, and interaction logs [Dix et al., 2004; Preece et al., 2007].

I used qualitative methods, including literature investigation, user observations, questionnaires and usability and accessibility evaluation and summarize these in section 3.4.2.

3.4 RESEARCH DESIGN AND METHODOLOGY USED IN THIS STUDY

In this section I describe the research process in my study in detail. Section 3.4.1 discusses my design research approach including the details of the two cycles of the design research phases. Then in section 3.4.2 I discuss the primary data collection methods. I summarize my use of the heuristic evaluation method, field evaluation and questionnaires to evaluate the usability and direct accessibility support in the DD. Section 3.4.3 summarizes the way in which I analyzed the primary data collected in this study and section 3.4.4 discusses the manner in which I addressed ethical issues.

3.4.1 Research Paradigm Used – Design Research

I used the design research as the primary research paradigm and my study involved two cycles of the design research phases. The first cycle encompasses the whole research process (section 3.4.1.1), while a second (inner) cycle developed the heuristics for evaluating the DD (sections 3.4.1.2).

3.4.1.1 Outer Cycle of the Design Research Phases

As discussed in section 3.2.1, design research involves analyzing the performance of a designed artefact or a system. In my study, this involved evaluating the usability and direct accessibility support of a selection of interfaces and educational game applications in the DD. Usability evaluation is an essential aspect of HCI to determine whether the target user groups will be able to use the given system or application to perform real tasks [Dix et al., 2004]. In the context of my study, the five phases of the outer design research cycle, as depicted in Figure 3.2, are:

1. *Awareness of problem:* The first phase of the design research process in my study involved realizing that the usability and accessibility of the applications installed on an interactive computer system had never been evaluated since its inception in 2002. Without evaluation, it is difficult to determine whether users will be able to use a given application to complete real tasks.

Thus, the outcome of this phase was recognizing the need to investigate the methods, documented in the literature, for evaluating the usability and accessibility to determine which of these methods are appropriate for the DD context.

2. *Suggestions:* This phase involved an extensive literature investigation of the various usability evaluation methods (Chapter 4), and the techniques for evaluating accessibility (Chapter 5). The choice of evaluation methods is typically influenced by eight factors (detailed in section 4.2.4). The relevant factors that affected the selection of evaluation methods for the DD are the following:

   - *The stage at which evaluation is done:* As stated in section 1.2, the DD has been a fully operational system since the first deployment in 2002. My study is thus a typical summative evaluation. The heuristic evaluation method is a relatively easy and effective method that is suitable for both formative and summative evaluation, provided appropriate evaluation heuristics are used in the evaluation. One of the central activities of this study was the derivation of evaluation heuristics for the DD. Further, the heuristic evaluation method is one of the techniques recommended by Henry [2007] for evaluating interactive systems' accessibility. This method provided a straightforward assessment of the level of direct accessibility support in the DD given that the DD does not support the use of assistive devices so users with disabilities cannot be included as study participants.
   - *Style of evaluation*: Evaluation can be done in controlled environments such as a usability laboratory or a natural environment of system use. However it was not feasible to use the controlled usability testing available in UNISA's well-equipped usability laboratory. In addition, the laboratory's observation and logging software is compatible only with applications running on the Microsoft Windows operating system, while applications on the DD run on the Ubuntu Linux operating system.

     Evaluation methods involving typical users can reveal usability problems that may be overlooked in heuristic evaluation. Thus, I also used direct observations at one of the centres where the DD is deployed in a field usability evaluation and used this in triangulating insights.
   - *Resource Requirements:* Every evaluation method is shaped by the availability of resources, including equipment, expertise, and participants' availability. The availability of expert usability/accessibility evaluators made the heuristic evaluation method appealing, since expert evaluators typically provide better results than training non-usability experts to conduct heuristic evaluation [Nielsen, 1992, 1994b].

     My main consideration in the field usability evaluation involved deciding the appropriate time to conduct evaluation to limit disruption to learning activities. Thus, evaluation sessions took place in the afternoons after the official school hours.

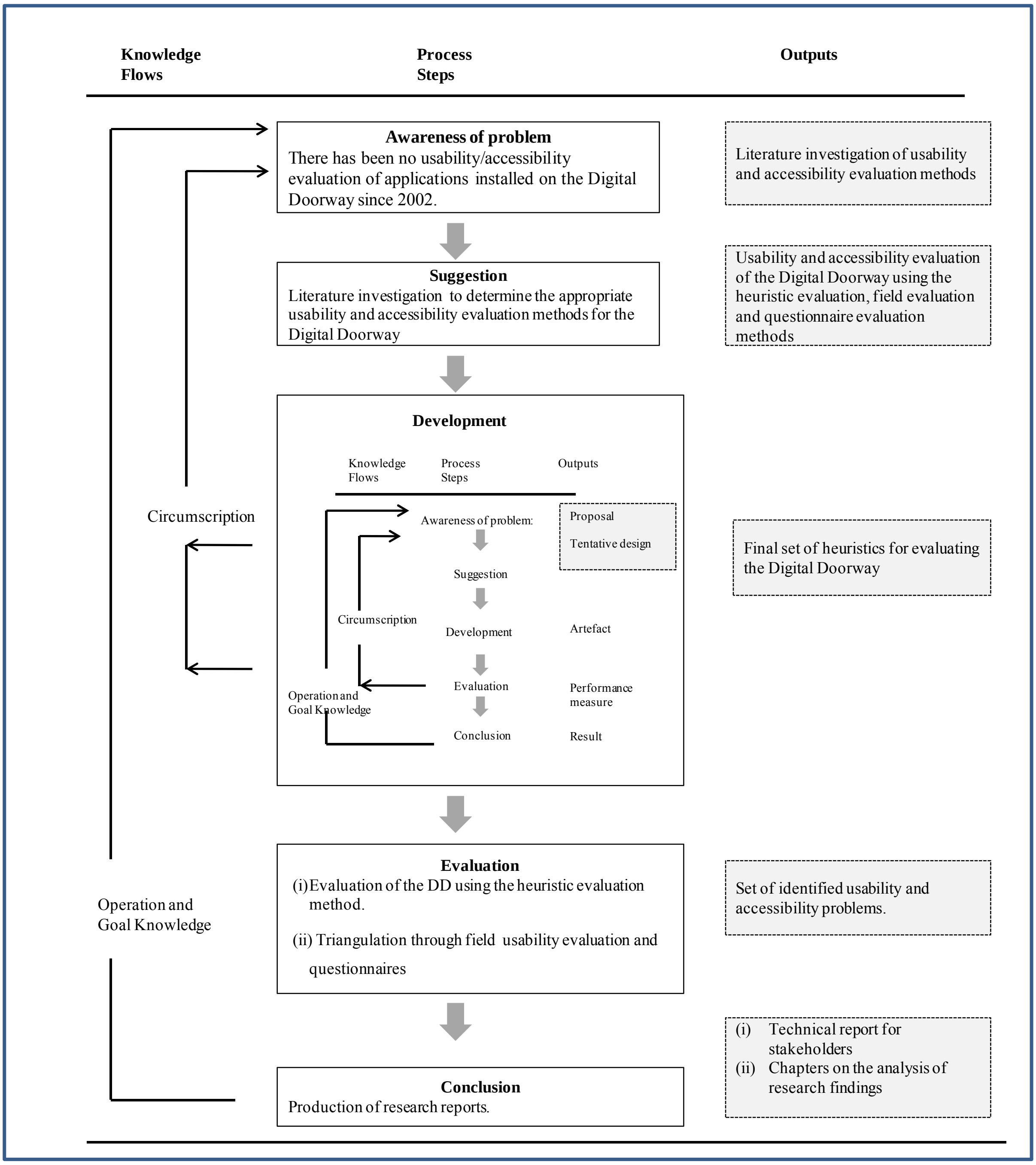


Figure 3.2: Outer cycle of design research phases

3. *Development:* Since the artefact in this study was not a physical object, the creation phase involved developing appropriate evaluation heuristics for the DD. This phase triggered the second (inner) design research cycle detailed in section 3.4.1.2. The development phase resulted in a set of multi-category evaluation heuristics that could be used to evaluate the applications installed on the DD, as presented in Chapter 7.

4. *Evaluation:* In this phase the DD was evaluated using the multi-category evaluation heuristics. The primary method for evaluating the DD was therefore a heuristic evaluation method by usability and/or accessibility experts as described in section 8.3.1.

I complemented heuristic evaluation with a field usability evaluation, detailed in section 9.2.1. This also provided a method for triangulating the usability and direct accessibility problems that emerged from the heuristic evaluation process.

In addition, I converted a selection of heuristics, which emerged from the development phase, into questionnaires to assess the usability of the DD. The questionnaires were used by participants in the field usability evaluation to assess the usability of the DD as described in section 9.3.

5. *Conclusion:* For the final phase of the design research paradigm, I generated two types of research reports. The first one was a technical report, followed by oral presentation of the findings from the study to the relevant stakeholders at the Meraka Institute. The second type of report documented research findings in Chapters 8 and 9.

3.4.1.2 Inner Cycle of the Design Research Phases

The inner cycle of the design research was triggered during the development phase of the outer design research cycle. It entailed the development of a set of multi-category heuristics for evaluating the DD. The five phases of this design research cycle, shown in Figure 3.3, are as follows:

1. *Awareness of problem:* The suggestion phase of the outer cycle of the design research process identified heuristic evaluation method as an appropriate method for assessing the usability and direct accessibility support provided in the DD. However, the value of the heuristic evaluation method lies in using appropriate heuristics for the specific system to be evaluated. Usability design principles in the literature, such as Dix et al. [2004] and Nielsen [1994b], focus on interface usability and cannot be used to evaluate accessibility. Likewise, the usability of educational game applications may not be adequately covered by general usability guidelines. Thus, for the interfaces and applications that were evaluated in this study, it was imperative to develop an integrated set of evaluation heuristics that address: general usability, direct accessibility and educational game usability.
2. *Suggestions:* This phase involved an extensive literature investigation of existing principles and guidelines for the design of usable and accessible interactive systems. This includes guidelines for the design of computer-based educational games (see Chapters 4, 5 and 6 for detailed discussion of the principles and guidelines examined).
3. *Development:* This phase involved the development of a multi-category evaluation heuristics for the DD. The heuristics were based on the general usability design principles, guidelines and heuristics (see section 4.3); the form interface design guidelines (section 4.4.1.1); the guidelines for the design of accessible interactive systems (section 5.5); and educational game design guidelines (section 6.5).

   To formulate the heuristics I examined the appropriateness of the principles and guidelines and accounted for the types of interfaces and applications to be evaluated, the intended types of users of the DD and the typical DD usage environment. Since the DD is intended as ‘walk-up-and-use’ system, users should be able to learn its use without external assistance.

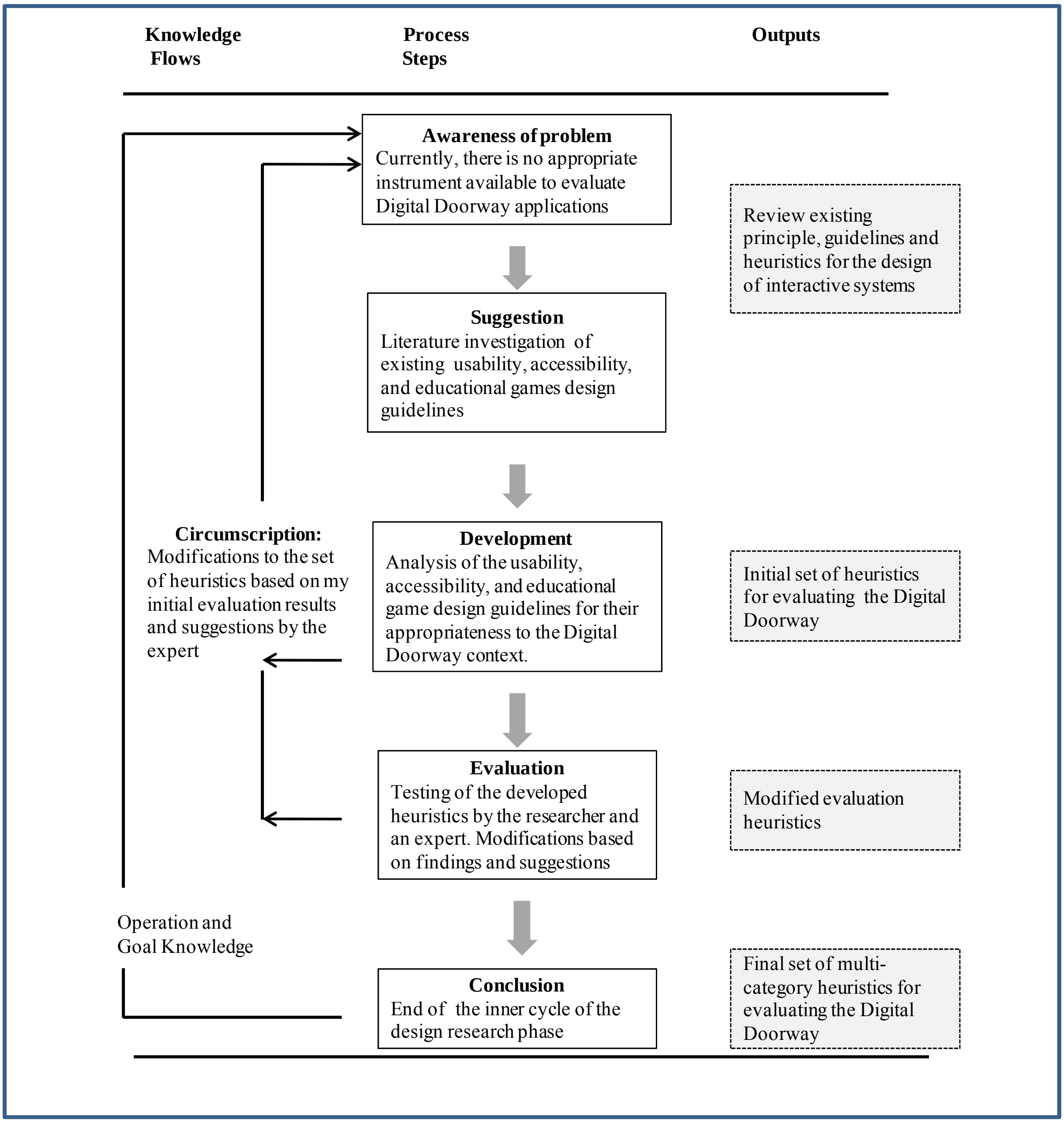


Figure 3.3: Inner cycle of design research phases

Following an initial literature investigation, the principles and guidelines were categorized according to the common issues they were meant to address. For example, grouping together all the principles and guidelines that are related to the principle of feedback.

Principles and guidelines that were not relevant to the DD system were eliminated taking into account the context for their use. For example, through this process I found the principles of multithreading and task migratability by Dix et al. [2004] were not essential to the context of use of the DD (the DD only supports the execution of one task at a time). Likewise, I also found that the Web Content Accessibility Guidelines (WCAG 1.0) [1999] (see section 5.5.2) is mainly aimed at Web-based systems. However, a few of WCAG 1.0 are relevant to the DD context. For example, guidelines relating to the use of colours and multimodal information presentation are useful to evaluate the direct accessibility support provided in the DD. The result of this phase was an initial set of evaluation heuristics.

4. *Evaluation:* I conducted a heuristic evaluation to test the initial set of evaluation heuristics, generated in the development phase, on the selected interfaces and applications in the DD. This determined the level of coverage of the developed heuristics. I detected usability and direct accessibility-related problems, which were not covered by the initial evaluation heuristics. I then re-visited the categorized principles and guidelines to determine those that are aimed at the identified problems but had not yet been matched to an evaluation heuristic, in order to update them.

   An HCI expert, with prior experience in usability and accessibility issues, used the modified heuristics, arising from the first round of testing and modification to the evaluation heuristics, to evaluate the DD. This expert was not included in the team of five experts who participated in the formal heuristic evaluation of the DD. The expert's suggestions led to further modifications to the evaluation heuristics. Thus, this phase resulted in a set of multi-category heuristics for evaluating the DD.

   Following the evaluation by the team of experts, two additional problems were identified by the team of evaluators for which there were no applicable heuristics. The heuristics were then modified to cover these problems.

   A final set of 77 heuristics was generated over three iterations, as summarized:

   - I tested, and then modified the first version of the heuristics on the DD.
   - A usability and accessibility expert tested the heuristics on the DD to determine the level of their coverage. I further modified the heuristics based on the expert's findings and suggestions.
   - A team of expert evaluators identified two additional problems, which could not be matched to any of the heuristics, and I made minor modification to the heuristics.

5. *Conclusion:* This phase marked the end of the inner cycle of the design research process and the resumption of the evaluation phase of the outer cycle.

3.4.2 Primary Research and Data Collection Methods

The following sub-sections overview the four research methods used in this study, namely, literature investigation, expert heuristic evaluation, user observations, and questionnaires.

3.4.2.1 Literature Investigation

The value of the heuristic evaluation method depends on using appropriate heuristics. Application-specific heuristics can be developed through literature investigation of existing principles and guidelines (*research-based method*) or from an analysis of usability problems obtained from previous studies (*evaluation-based method*) [Paddison and Englefield, 2004]. The methods for developing application-specific heuristics are discussed in section 4.2.3.1.2.

In this study, heuristics for evaluating the DD were developed using the research-based method, which entailed determining the appropriateness of existing principles and guidelines reported in the literature. This largely focused on the principles and guidelines for designing usable and accessible interactive systems, and guidelines for the design of computer-based educational games. Sections 4.3, 4.4.1, 5.5, and 6.5 discuss in detail the design principles, guidelines and heuristics examined, as well as those that are applicable to the evaluation of the DD.

3.4.2.2 Expert Heuristic Evaluation

The heuristic evaluation technique, pioneered by Nielsen and Molich [Dix et al., 2004; Preece et al., 2007], involves expert evaluators independently critiquing an interface, at formative and formative stages, using a set of evaluation heuristics. The heuristic evaluation method identifies potential usability problems in a straightforward, flexible and cost effective manner. Section 4.2.3.1 discusses the method in detail, including requirements for application-specific heuristics and multiple evaluators to ensure effective evaluation. As discussed in section 8.3.1, my study involved five usability and/or accessibility experts evaluating a selection of interfaces and applications installed on the DD using the developed heuristics.

3.4.2.3 User Observations

As section 4.2.3.6 describes in detail, observations in HCI research involve observing users interacting with a system [Dumas, 2003; Millen, 2000; Preece et al., 2007], in either controlled or field environments [Dumas, 2003; Preece et al., 2007]. The data gathered enables researchers to construct a rich understanding of an application or system's context of use [Preece et al., 2007]. As discussed in detail in section 9.2.1, I observed DD usage at a local secondary school with the primary objective of triangulating the data obtained from the heuristic evaluation process.

3.4.2.4 User Questionnaires

Questionnaires have become well established in HCI research because they can, potentially, collect data on users' attitudes, preferences and product evaluation from a large numbers of users with relative ease [Dix et al., 2004; Ozok, 2008; Preece et al., 2007]. Section 4.2.3.5 describes in detail the use of questionnaires in HCI research. They are used as a stand-alone usability evaluation method or combined with other evaluation methods [Dumas, 2003]. Further, questionnaires can be variously structured. The types of responses can be fixed or open ended, where respondents provide as much or as little information as they prefer [Ozok, 2008; Preece et al., 2007]

Questionnaires were used as a complementary evaluation method to evaluate the DD in this study and their results were triangulated with data from the field usability evaluation. As section 9.3.1 discusses, questions in the questionnaires were fixed, with the opportunity for respondents to include open ended comments.

3.4.3 Data Analysis

This section provides a summary of how the primary data generated in my study was analyzed.

3.4.3.1 Heuristic Evaluation Data Analysis

Following the heuristic evaluation by all experts, I compiled an aggregate of usability and direct accessibility-related problems identified by consolidating those that were related and of similar nature. Identified usability problems were analyzed according to the number of evaluators that detected the specific problems. I also categorized the set of identified problems according to the evaluation heuristics that had been violated by the specific problem. Further analysis was based on the specific interface/application in which the problems were located. Descriptive statistics also provided summaries of the complete set of unique usability problems identified as well as the number of problems identified by each evaluator. A detailed discussion of the heuristic evaluation data analysis is provided in sections 8.3.2 to 8.3.5.

3.4.3.2 Field Usability Evaluation Data Analysis

I documented observational data in handwritten field notes and video recordings of sessions. This data was then analyzed by first categorizing it according to the application used by participants, that is, *What-What Mzansi*, *OpenSpell* and *Themba's Journey*. Similar usability and/or accessibility problems were then grouped together before using descriptive texts to notate them. Finally, I compared the set of usability and direct accessibility-related problems detected by the expert evaluators with the actual problems experienced by participants in the field usability evaluation. A detailed discussion of the field evaluation data analysis is in sections 9.2.2 and 9.2.3.

3.4.3.3 Questionnaire Data Analysis

Questionnaires returned by the participants were analyzed by summarizing the participants' ratings to specific statements in the questionnaires. I read open-ended comments provided by participants, taking note of those that related to similar problems. The comments were recorded verbatim as described by the participants. The ratings and comments of participants to the questionnaires were compared with participants' behaviour during the evaluation sessions to verify and clarify the ratings and comments. Sections 9.3.2 and 9.3.3 provide detailed discussion of the analysis of the questionnaires.

3.4.4 Ethical Considerations

According to Resnik [2007], ethics can be loosely defined as the norms for conduct or behaviour that distinguish between what is acceptable and what is not. Research ethics policies are typically formulated by professional bodies, government agencies and academic institutions to guide research conduct. In this study ethical considerations were addressed on three levels:

1. This study complies with the UNISA research ethics policy [2007]. Ethical clearance was obtained from the CSET Research and Ethics Committee (Appendix A).

2. Expert evaluators were provided with an informed consent document to sign (Appendix B). The evaluators were also assured of their anonymity and the confidentiality of information they provide.
3. Because the field evaluation participants were minors under the age of eighteen, formal permission was sought (Appendix C) and obtained from the school principal to conduct the study at the school. In addition, informed consent was obtained from parents/guardians of participants who volunteered to take part in the study (Appendix D). Participants were also assured that video recordings of evaluation sessions will be protected and none of the images will be used in publications.

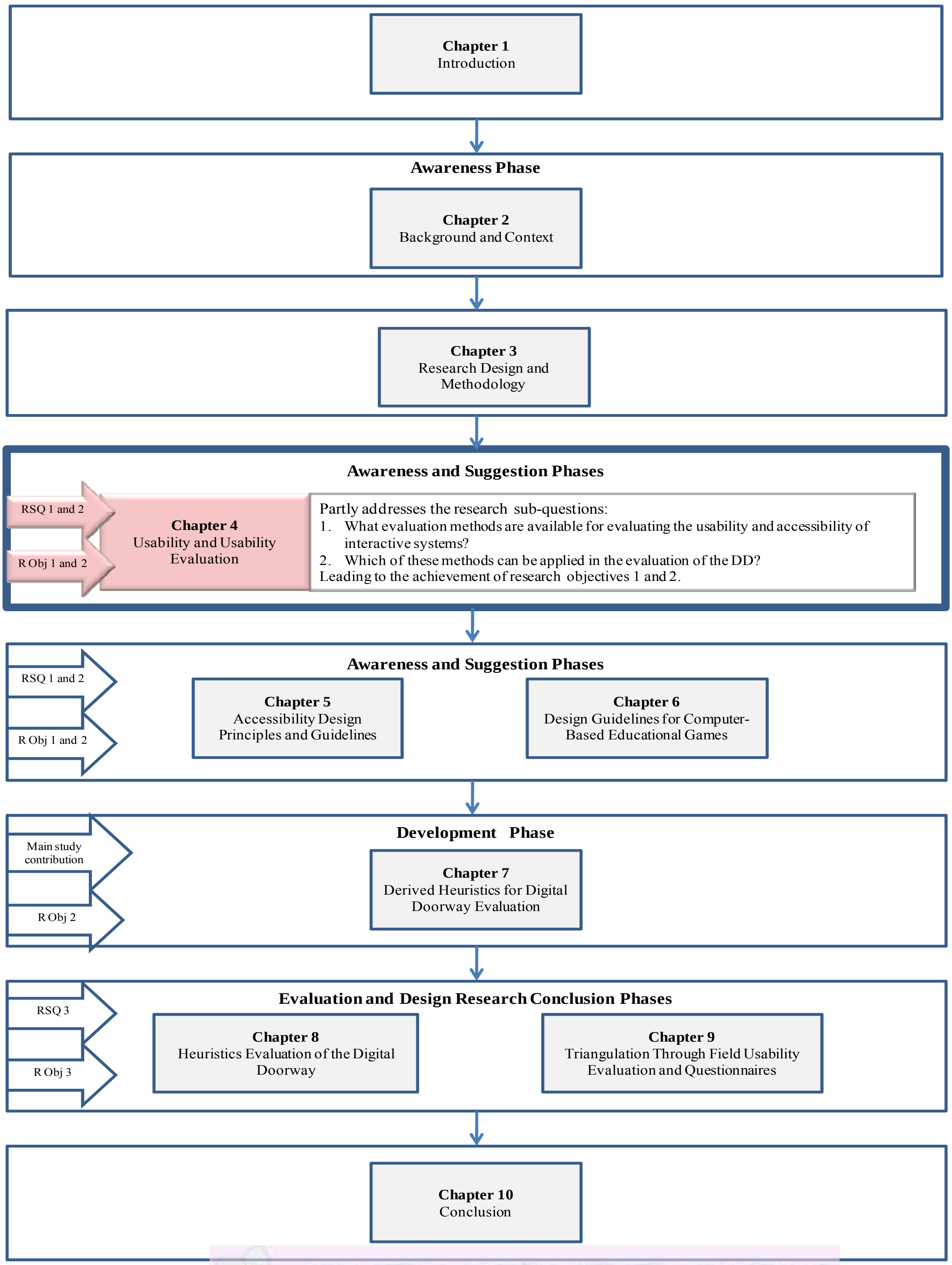
3.5 CONCLUSION

In this chapter, I overviewed different types of research paradigms together with their underlying philosophies, and the specific research design and methodology in the study. I detailed the application of two cycles of the five phases of design research paradigm and introduced the primary data collection and analysis methods. I also summarized measures taken to conform to the UNISA research ethics policy.

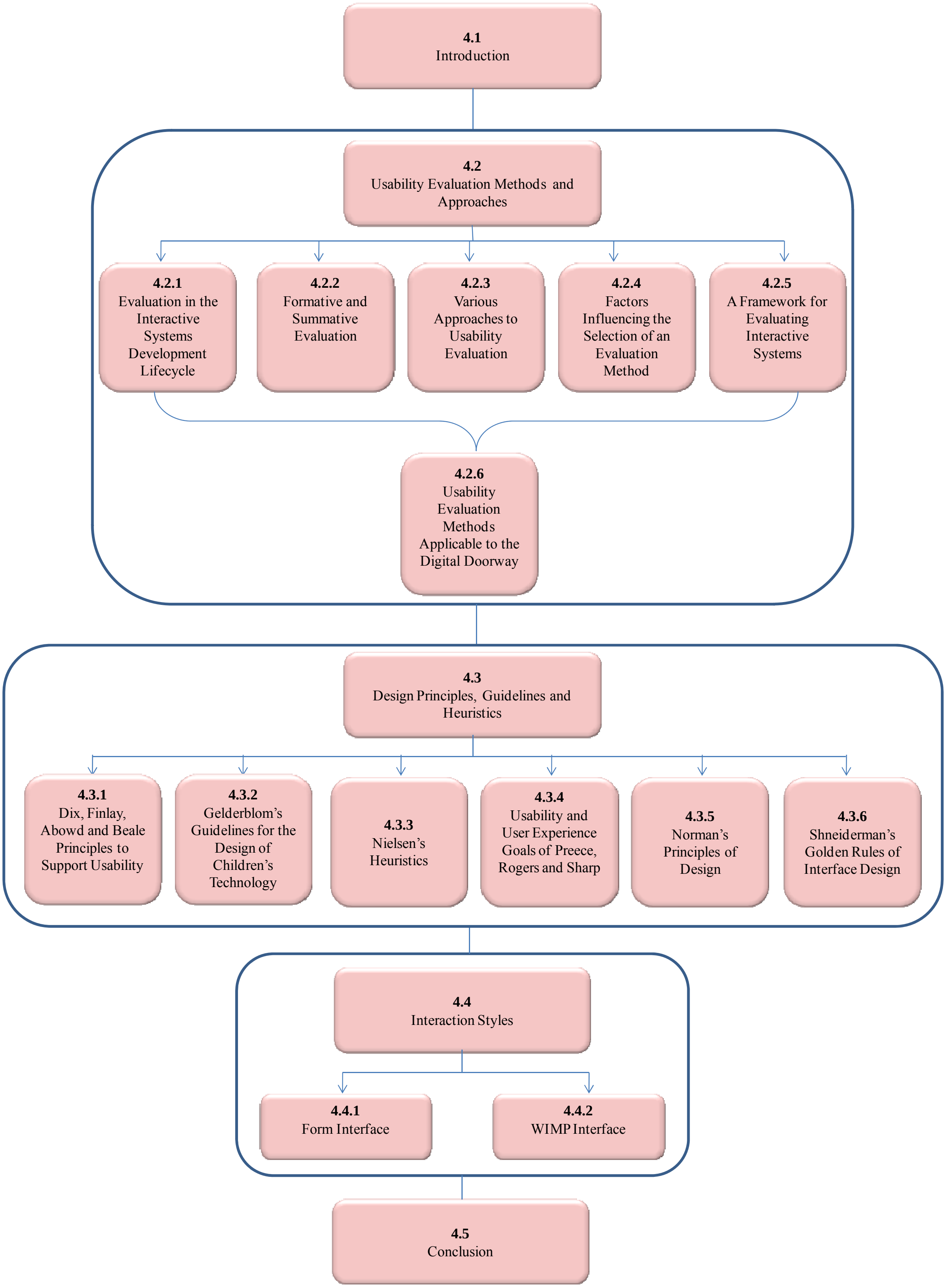
In the next three chapters, I discuss the literature on existing guidelines for the design of usable and accessible systems and examine the methods available for evaluating the usability and accessibility of interactive systems.

CHAPTER 4: USABILITY AND USABILITY EVALUATION

The stage of Chapter 4 in the dissertation



Map of Chapter 4



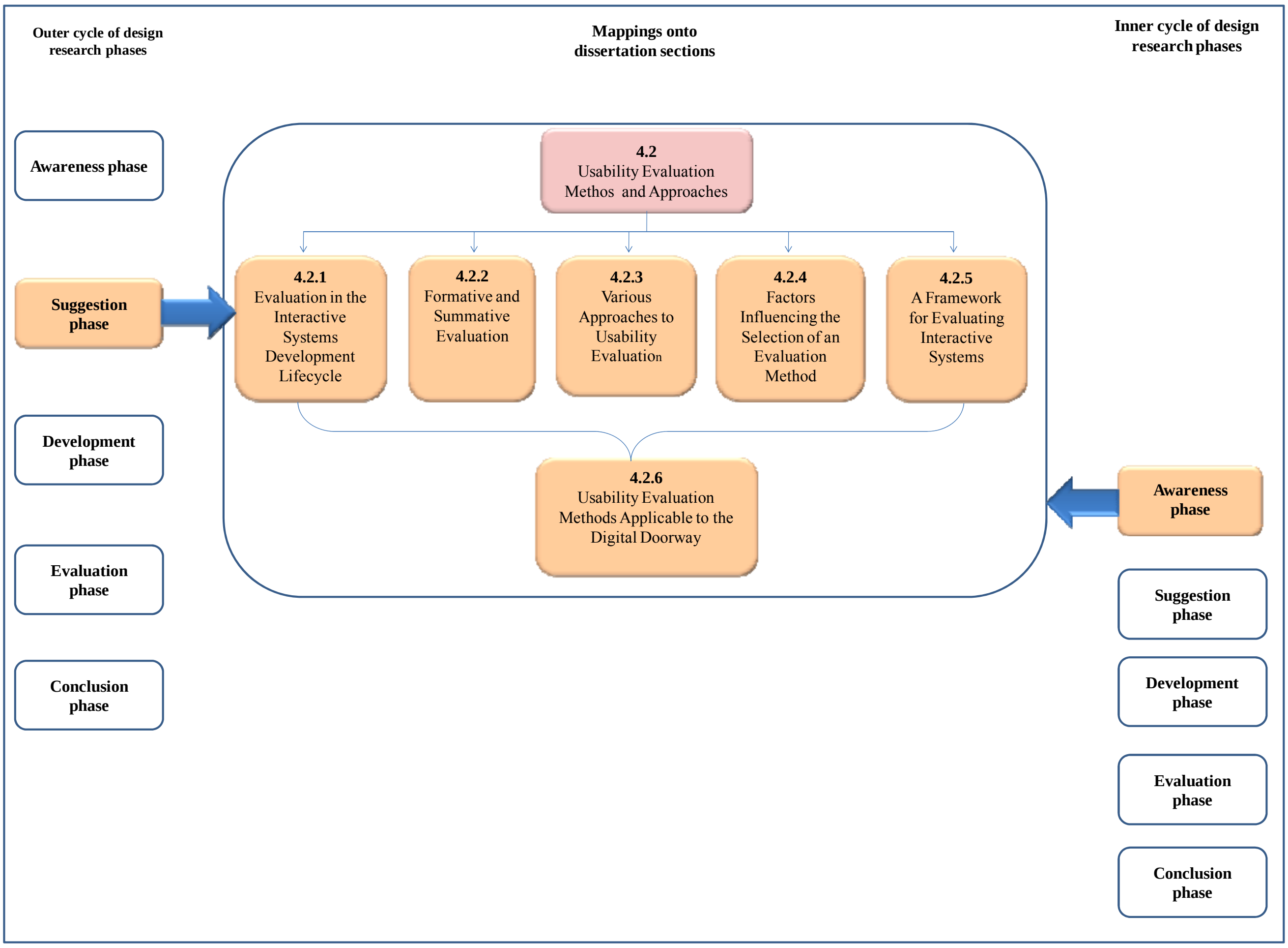
4.1 INTRODUCTION

This chapter is the first of three literature investigation chapters that represent the suggestion phase for the outer design research cycle, and the awareness and suggestion phases of the inner cycle of my design research (discussed in section 3.4). It partly addresses two of my research sub-questions: sub-question 1, What evaluation methods are available for evaluating the usability and accessibility of interactive systems? and, sub-question 2, Which of these methods can be applied in the evaluation of the DD?

I start by reviewing seven different types of usability evaluation methods (section 4.2) to determine which methods might be suitable to evaluate the selected applications and interfaces of the DD. In section 4.3, I look at existing principles, guidelines and heuristics for the design of usable interactive systems to assess their appropriateness in the development of evaluation heuristics for the DD. Section 4.4 provides an overview of the two main interface styles in the DD, with detailed discussion of the guidelines for the design of form interfaces. Section 4.5 concludes the chapter.

4.2 USABILITY EVALUATION METHODS AND APPROACHES

Mapping of section 4.2 to the design research phases



This section maps onto the suggestion phase of the outer cycle of the design research, and the awareness phase of the inner cycle of the design research. For the suggestion phase the methods for evaluating the usability of interactive systems are examined to determine the most appropriate methods for evaluating the applications and interfaces on the DD. The awareness phase highlights the need for application-specific heuristics for the DD.

Evaluation can be defined as "the process of systematically collecting data that informs us about what it is like for a particular user or group of users to use a product for a particular task in a certain type of environment" Preece et al. [2002:317]. This section discusses the place of evaluation in the interactive system development lifecycle and seven types of usability evaluation methods; and thus partly answers my research sub-question 1: What evaluation methods are available for evaluating the usability and accessibility of interactive systems?

4.2.1 Evaluation in the Interactive Systems Development Lifecycle

The traditional waterfall lifecycle model represents system development activities in phases that are carried out linearly, where a previous phase must be completed before the next one can begin. This model provides for limited feedback to earlier phases but evaluation was not built into it [Preece et al., 2007]. Rather, evaluation is an iterative and ongoing process that should be carried out throughout the design process and not a single phase that comes at the end of development. Evaluation should start early, and the results of evaluation feed back into the design. The longer evaluation is delayed, the more difficult and costly it becomes to correct problems uncovered at later stages.

One model that highlights the central role of evaluation is the *Star lifecycle model* (Figure 4.1), developed by Hartson and Hix [1989] (as cited by Costabile [2001] and Preece et al. [2007]). Instead of enforcing an orderly, top-down approach to development, this model provides developers with the flexibility of moving from one activity to another in any order. Activities in the model are highly interconnected, with evaluation at the center of all activities. Following the completion of each activity, its result must be evaluated. Thus a development project may start from any point in the star, for example the evaluation of an existing product, and followed by any stage. This enables the design to evolve gradually.

4.2.2 Formative and Summative Evaluation

Evaluation methods can be grouped according to their role in the development lifecycle. *Formative evaluations* capture user requirements for a product and guide the decision between design alternatives and are thus carried out during the development stage. Typically, formative evaluations are iterative and explorative in nature, with feedback from the evaluation used to modify and refine the design [Kotzé and Johnson, 2004; Preece et al., 2007].

Summative evaluations are performed to assess a finished or existing product and are used by developers to demonstrate to clients that the finished product meets their requirements.

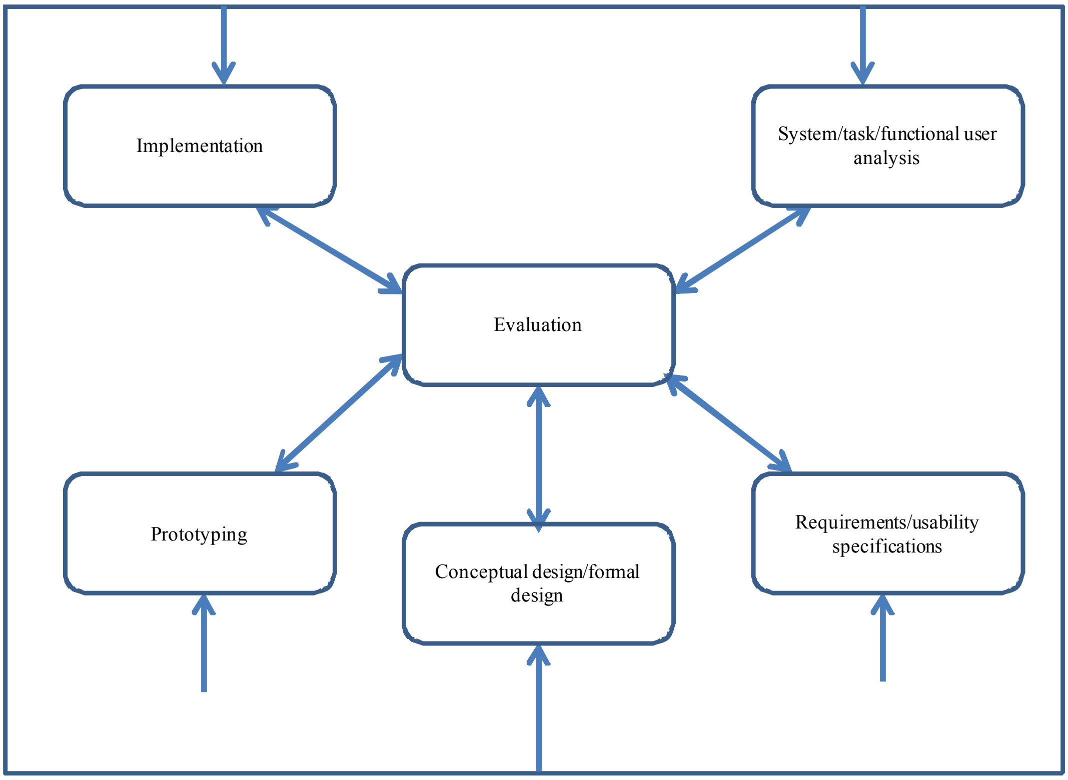


Figure 4.1: Hartson and Hix's star lifecycle model (as depicted in Preece et al. [2007])

Summative evaluation usually focuses on one or two specific issues, for instance usability, and can be used to identify any residual problem in the product before deployment [Kotzé and Johnson, 2004; Preece et al., 2007].

4.2.3 Various Approaches to Usability Evaluation

An application' usability can be evaluated through various evaluation methods. Often the methods are generally classified according to their involvement of expert analysts or end users [Dix et al., 2004]. Typical usability evaluations through expert analysis are the heuristic, cognitive walkthrough and model-based evaluations. Evaluations involving user participation include query techniques, such as interviews and questionnaires, observations and experimental evaluation.

Preece et al. [2007] used a different classification scheme to identify three approaches to evaluation: usability testing, field studies and analytical evaluation. Each of these approaches can use several methods, including user observations, interviews, questionnaires, expert analysis, and user testing. Although the authors used different classification schemes/terminologies, the approaches are quite similar and not mutually exclusive in their implementation. In a typical evaluation process, multiple evaluation methods/approaches are employed to get multiple perspectives on the focus of the evaluation.

In the following sub-sections I review seven different methods for evaluating the usability of interactive systems to determine the most appropriate methods for evaluating the selected applications and interfaces of the DD.

4.2.3.1 Heuristic Evaluation

4.2.3.1.1 Introduction

In heuristic evaluation experts assess whether an artefact conforms to a set of usability guidelines [Dix et al., 2004; Nielsen, 1994b; Preece et al., 2007]. A heuristic can be defined as a guideline, general principle or rule of thumb. It can guide a design decision in formative evaluation where a design is in early development; or it can be used to critique a decision that has already been made in summative evaluation of a fully functioning system [Dix et al., 2004].

The heuristic evaluation method was originally developed by Nielsen and Molich [Dix et al., 2004; Preece et al., 2007]. It is a flexible and relatively inexpensive method that does not require user involvement, which is why it is sometimes called a *discount usability* method [Dix et al., 2004]. However, its real value lies in the use of appropriate evaluation heuristics [Jeffries, Miller, Wharton and Uyeda, 1991; Preece et al., 2007]. Nielsen's heuristics (section 4.3.3) for example, are sometimes considered to be too general, necessitating the development of application-specific heuristics.

4.2.3.1.2 Methods for Deriving Application-Specific Heuristics

The original set of heuristics developed by Nielsen [1994b] while suited to single-user, desktop-based and task-oriented applications are inadequate for evaluating the usability of ever increasing ubiquitous, groupware applications. For instance, compare designing traditional transaction processing applications that are easy to use, minimize user error and support rapid task completion with designing game applications that are pleasurable but challenging to the user [Korhonen, 2010]. To evaluate these usability characteristics will necessitate specific heuristics that focus on the varying goals.

Typically, researchers develop application-specific heuristics to address specific usability goals of different application domains. For example the heuristics for evaluating Web-based e-learning applications [Ssemugabi and De Villiers, 2007].

There are three approaches to developing application-specific heuristics (i) the *research-based method* (ii) *evaluation-based method* [Ling and Salvendy, 2005; Paddison and Englefield, 2004], and (iii) the *multi-method* approach that combines the first two approaches [Sim, Read and Cockton, 2009].

1. In the *research-based method*, application-specific heuristics are developed by consulting the literature to investigate the main characteristics of the application domain to determine the specific usability goals that must be addressed by the set of heuristics. This method can also derive application-specific heuristics by examining existing usability principles and guidelines [Paddison and Englefield, 2004]. Desurvire, Caplan and Toth [2004] derived heuristics to evaluate the playability of games based on a literature study of heuristics for playtesting of games and productivity applications. The effectiveness of the heuristics was tested by expert evaluators and the results compared with the problems identified through user testing.
2. In the *evaluation-based method*, application-specific heuristics are developed by analyzing the results from previous evaluation studies to identify the main problems that emerged. Berry [2003]

used this approach to develop heuristics for evaluating the usability of notification systems by analyzing their common usability problems and synthesizing the problems into eight categories of heuristics.

3. In a *multi-method* approach to developing application-specific heuristics, Sim et al. [2009] developed heuristics for evaluating the usability of computer-assisted assessment (CAA) applications. The method involved creating a collection of usability problems through student surveys (after CAA exams); heuristic evaluations, using Nielsen's [1994b] heuristics to determine its applicability to CAA domain; and a literature review to expand the collection of usability problems. The set of CAA-specific evaluation heuristics was derived from synthesizing these usability problems.

4.2.3.1.3 Validating the Effectiveness of Application-Specific Heuristics

The effectiveness of a heuristic set depends on its ability to adequately capture all the significant usability problems in an application. However, the total number of usability problems identified through the heuristics may not necessarily indicate the effectiveness of an application-specific heuristic set. As Sim et al. [2009] discusses, an effective heuristic depends on the following criteria:

- *Correctness* refers to the terminology used in describing the heuristics. This criterion is used to evaluate whether the description of the heuristics provide sufficient information to evaluators who will use the heuristics in the evaluation process.
- *Coverage and thoroughness* of the heuristic set refer to the extent to which the heuristics adequately cover the domain being evaluated.

Another method for validating the effectiveness of a set of application-specific heuristics involves *comparison of the results* obtained from the application of the derived heuristics in evaluation studies relative to those obtained by applying standard heuristics, such as, Nielsen's heuristics [Ling and Salvendy, 2005].

4.2.3.1.4 The Heuristic Evaluation Process

Heuristic evaluation involves evaluators examining an interface independently and judging its compliance with the set of heuristics. Its main goal is the identification of usability problems in the design so that they can be corrected as part of an iterative design process [Nielsen, 1992; Nielsen, 1994b].

It is important for each of the evaluators to independently evaluate the application to ensure unbiased assessment. Only after all the evaluators have completed their review should there be any form of interaction between them [Dix et al., 2004; Nielsen, 1994b]. The standard evaluation process is composed of three stages [Preece et al., 2007]:

1. In the first stage, the *briefing session*, evaluators are provided with necessary information regarding the evaluation. This is usually carried out by following a checklist to ensure that each of the evaluators receive the same briefing.

2. In the *evaluation period*, each evaluator independently goes through the interface at least twice. Usually, the first pass aims to familiarize evaluators and provide an overview of the flow of the interaction. The second pass focuses on specific interface components to identify potential usability problems.

   Evaluators or scribe record violations of any of the heuristics while examining the interface. In recording, the evaluator must be as specific as possible by noting all the problems associated with a particular element separately and indicating which of the heuristics have been violated. For instance, if five problems have been identified for a single element, then all the five problems should be recorded with reference to the specific heuristics that have been violated. The recording of each problem separately offers two main benefits. Firstly, if all the problems relating to a given component are identified, the probability of repeating the same problem in a new design will be reduced. Secondly, even if it is impossible to redesign the component or correct all the problems associated with it, it will at least be possible to correct some of the problems if they are all known [Nielsen, 1994b].

3. The last stage, the *debriefing session*, collects and documents the results of all the problems identified by all the evaluators in a report including the calculation of the mean severity ratings for each problem. This report should also indicate which heuristics were violated. Although not mandatory, the report may also list recommendations for correcting identified problems.

A variant of the heuristic evaluation process by Xerox Corporation [1996] consists of five sets of activities classified according to the person(s) responsible for their execution. These activities are [Xerox Corporation, 1996]:

1. *Getting ready* involves the project leader identifying the heuristics that will be used for the evaluation; selecting the people to perform the evaluation; scheduling the location; date and time for each evaluator; compiling documents required for the evaluation, such as user and system profile, user tasks/scenarios and problem report sheet; and, deciding whether each evaluator will evaluate the system individually, or as a group, and whether or not a scribe will be allocated to evaluators.

2. *System evaluation* involves evaluators experimenting with the system to obtain a feel of how it works before listing the components which violate any of the heuristics. This stage is similar to the second stage of evaluation process identified by Preece et al.[2007].

3. *Result analysis* involves the group activities which review the usability problems identified by all evaluators. An affinity diagram is constructed to group the identified usability problems together, with severity ratings allocated to them. The evaluators then provide recommendations for correcting the problems.

4. *Reporting the result* involves the team leader compiling the results from the third stage. Here, the team leader lists, in a format that is appropriate to report’s audience, all the identified usability problems; the heuristics violated; the severity ratings allocated; and recommended corrections.

5. *Debriefing* involves the team leader making an oral presentation of the evaluation report. In addition to highlighting the major usability problems identified, the positive aspects of the system are also noted.

4.2.3.1.5 How Many Evaluators will Provide an Optimal Result?

All the problems identified by multiple evaluators are aggregated because it is difficult for a single evaluator to identify all the usability problems in a design or an application. Nielsen [1994b] analyzed the proportion of usability problems found in an interface by varying number of evaluators and found that, averaged over six projects, a single evaluator can identify just about 35 percent of usability problems (see Figure 4.2). The proportion of problems identified increases with the number evaluators included, such that about five evaluators can identify an aggregate of 75 percent usability problems. However, after this optimal level, adding more evaluators does not significantly increase the number of problems identified; so ten evaluators identify about 85 percent problems and fifteen uncover ninety percent problems. The actual number of evaluators depends on the type of application being evaluated. Where increased usability is required, the use of more evaluators is beneficial. However, any evaluation process should be conducted with at least three evaluators [Nielsen, 1994b].

4.2.3.1.6 Assigning Severity Ratings to Usability Problems

Severity ratings for usability problems provide an estimate of the degree to which identified problems will negatively affect the application's usability should it be deployed. In addition, severity ratings can guide allocating scarce resources to the most serious usability problems [Nielsen, 1994b].The level of severity of identified problems is influenced by three factors, namely [Nielsen, 1994b]:

1. *The frequency of the problem:* This relates to the number of times the problem occurs.
2. *The impact of the problem:* This refers to the level of ease with which users are able to recover from the problem should it occur.

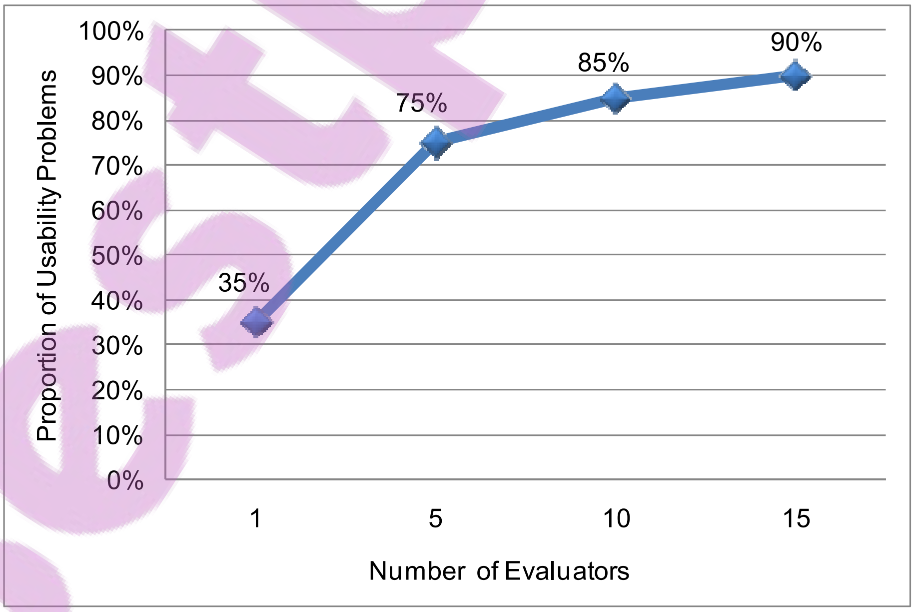


Figure 4.2: Curve showing proportion of usability problems found in heuristic evaluation using various numbers of evaluators [Nielsen, 1994b]

3. *The persistence of the problem:* This refers to whether the problem will be a once-off occurrence that users can recover from once they know about it or if users will be confronted with the same problem several times.

Nielsen [1994b] recommends that the severity of usability problems found in a given interface be ranked on a five-point scale using the following scheme:

1. 0 = I don't agree that this is a usability problem at all.
2. 1 = Cosmetic problem only, need not be fixed unless extra time is available on project.
3. 2 = Minor usability problem, fixing this should be given low priority.
4. 3 = Major usability problem, important to fix, so it should be given high priority.
5. 4 = Usability catastrophe, imperative to fix this before product can be released.

Wilson and Coyne [2001] argue that organizations should apply a common severity rating scale for usability and non-usability problems to ensure that adequate resources are committed to correcting the usability problems in the same way as other non-usability problems. For example, an organization that uses a four-point scale to rank the severity of programming problems should also use the same scale for usability problems, as illustrated in Table 4.1.

4.2.3.1.7 The Effect of the Evaluator on the Heuristic Evaluation Process

Various authors, such as Hornbaek and Frojaer [2008] and Hertzum and Jacobsen [2003], use the term, *evaluator effect*, to describe the variations in the nature of usability problems found by expert evaluators, as well as the severity ratings assigned to identified problems. In addition to the subjective nature of expert heuristic evaluation, other factors that contribute to evaluator effect are [Hertzum and Jacobsen, 2003]:

- *Vague goal analysis* can lead to variances in the kind of problems detected by evaluators. If a

Table 4.1: Four-point severity rating scale [Wilson and Coyne, 2001]

Severity Ratings	**Description**
1 – Severe	A catastrophic usability problem that causes system failure or unrecoverable loss of data. A usability bug that is likely to cause frequent data integrity errors. These problems have no workaround.
2 – High	A serious condition that impairs the operation, or continued operation, of one or more product functions and cannot be easily circumvented or avoided. The software does not prevent the user from making a serious mistake. The usability problem is frequent, persistent and affects many users. Standards are seriously violated.
3 – Medium	A non-critical, limited problem (no data lost or system failure). It does not hinder operation and can be temporarily circumvented or avoided. The problem causes users moderate confusion or irritation.
4 – Low	Non-critical problems or general questions about the product. There are minor inconsistencies that cause hesitation or small aesthetic issues like labels and fields that are not aligned properly.

specification of which elements of a given interface are to be assessed is absent or without clarity, evaluators can simply pick those they deem worthy of being evaluated.

- *Vague evaluation procedure* can lead to inconsistency between evaluators in using the evaluation heuristics. Also, the level of an evaluator's experience can also affect the way in which the heuristic evaluation is performed.
- *Vague problem criteria* also contribute to variations among problems detected by expert evaluators due to differences in their understanding of the system or application being evaluated. Criteria that specify and shared the definition of what constitute usability problems can significantly reduce evaluator effect; but may result in evaluators overlooking or missing usability problems.

4.2.3.1.8 Benefits of Heuristic Evaluation

In spite of the evaluator effect problem characteristic of heuristic evaluation, the method is accepted as a straightforward and cost-effective way to improve the usability of interactive systems' interfaces and has gained widespread use. The method is effective in identifying large numbers of potential usability problems, both major and minor [Nielsen, 1994b]. It is a flexible method that can be used on prototypes and fully functioning systems. The suitability of the method at design stage means that the cost of correcting identified problems will not be as high as when problems are identified at a later stage. Heuristic evaluation does not require the use of expensive and sophisticated equipment as is the case with other methods, such as, laboratory usability testing [Dix et al., 2004]. In addition, fewer practical and ethical issues are involved in heuristic evaluation relative to methods in which users participate in the evaluation [Preece et al., 2007].

4.2.3.1.9 Limitations of Heuristic Evaluation

Despite the benefits of the heuristic evaluation method, a number of shortcomings are associated with the method. Firstly, to be effective, the set of heuristics must be appropriate for the specific application to be evaluated. Also, multiple evaluators, preferably those with usability experience are required. Finally, evaluators may overlook problems that might impact on real users [Jeffries and Desurvire, 1992; Jeffries et al., 1991; Nielsen, 1994b].

4.2.3.1.10 Conclusion to Heuristic Evaluation

Heuristic evaluation is an effective method for identifying both minor and major problems in a user interface. The method is inexpensive and easy to learn. Its flexibility makes it possible to adapt the method to a specific project's needs and circumstances. To improve the effectiveness of the method, a minimum of three evaluators should be used.

4.2.3.2 Cognitive Walkthrough

4.2.3.2.1 Introduction

Cognitive walkthrough is derived from cognitive science theory such that expert evaluators, with cognitive theory skills [Wharton, Braffort, Jeffries and Franzke, 1992], inspect an application by stepping through a set of tasks. Cognitive walkthrough aims to assess the learnability of systems where users learn to use the system by exploration rather than reading through user manuals [Cockton, Lavery and Woolrych, 2008; Dix et al., 2004; Preece et al., 2007; Wharton et al., 1992]. The method is similar to the code walkthrough used in software engineering, and does not require user participation. Cognitive walkthrough is relatively inexpensive and flexible. It can be used for formative and summative evaluation before user testing [Wharton et al., 1992].

4.2.3.2.2 The Cognitive Walkthrough Process

To perform cognitive walkthrough, evaluators require the following information [Dix et al., 2004]:

- A specification or prototype of the system under review which should be detailed enough to permit effective review.
- A description of typical tasks that users will perform on the system.
- The sequence of actions which users need to execute in order to complete the tasks.
- The characteristics of typical users of the system.

Once the specified information has been provided, evaluators then walk through the action sequences to identify potential usability problems. In doing so, they attempt to answer the following questions [Dix et al., 2004]:

- Does a particular action enable the user to achieve his/her goal? For example, if the goal of the user is to copy a file from one location to another, will the action result in that effect?
- Are the permissible actions visible to the user? For example, can the user see the menu item required for the action?
- Having seen the correct item, will the user be able to recognize that it is the one s/he is looking for?
- Will the user be able to determine if s/he has successfully achieved his/her goal from the feedback provided by the system?

4.2.3.2.3 Documenting the Cognitive Walkthrough Process

Cognitive walkthrough is document-driven, requiring adequate record of both the ‘good’ and the ‘bad’ in the design or system being evaluated. Various authors recommend standardized evaluation forms for recording the evaluation, including [Dix et al., 2004; Preece et al., 2007]:

- A cover form, which records the date, time, names of evaluators, and the materials for the walkthrough.
- A separate form for each of the action sequences which records the answers to the questions outline in section 4.2.3.2.2.
- A special usability problem report sheet, which records any negative answer to the questions outline in section 4.2.3.2.2. The report also includes a detailed description of the usability problem, the severity of the problem, its frequency, and implication for users.

The documentations and reports generated from the cognitive walkthrough process should assist designers in prioritizing corrections to the design, as it may be infeasible to correct all the problems identified.

4.2.3.2.4 Benefits and Limitations of Cognitive Walkthrough

The flexibility of cognitive walkthrough makes it appropriate for formative and summative evaluation [Wharton et al., 1992]. In addition to predicting potential usability problems, expert evaluators also suggests corrections or modifications to the system [Liu, Osvalder and Dahlman, 2005]. The cognitive walkthrough is disadvantaged by the assumption that the evaluator has cognitive theory skills. Also, it does not address other measures of usability like the application efficiency [Wharton et al., 1992].

4.2.3.2.5 Conclusion to Cognitive Walkthrough

Cognitive walkthrough is an effective method for identifying certain usability problems early in the development process. It is a useful user-centred approach to development without requiring user participation. However, the method can be laborious and time-consuming. Also, other measures of application usability might be overlooked [Wharton et al., 1992].

4.2.3.3 Model-Based Evaluation

4.2.3.3.1 Introduction

Model-based evaluation methods are used by experts to predict user performance in systems without requiring user involvement. Expert evaluators attempt to predict usability measures through formulas or simulations of models [Kieras, 2008; Preece et al., 2007]. The technique aims to obtain some usability results before a given design or prototype is implemented or tested with human participants. The initial development of predictive models can be time-consuming. However, subsequently the models allow rapidly predicting usability by calculation or running a simulation. They also enable swiftly exploring the effect of modifications to the design by making small changes to the model, thus making the design-evaluate-redesign process faster [Kieras, 2008]. Two well-known model-based evaluation methods, the goal, operator, method and selection rule (GOMS) and the keystroke-level models are described below.

4.2.3.3.2 The GOMS Model

The Goals, Operators, Methods and Selection rules, commonly know as the GOMS model predicts an expert user's performance in a computer-based task in terms of the user's goals and the selection of methods required to achieve them [Dix et al., 2004; Kieras, 2008; Preece et al., 2007]. The GOMS model has the following four elements [Dix et al., 2004; Kieras, 2008; Preece et al., 2007]:

1. *Goals:* Represent what the user wants to achieve; for example, to print a document.
2. *Operators:* Actions the user needs to perform to achieve his/her goals. The actions may involve both cognitive processing and physical activity. For example, the user may need to recall where in the menu palette the print option is located. Furthermore, the user may need to physically position the cursor on the appropriate menu to make a selection.

3. *Methods:* Refer to the various ways in which the user can achieve his/her goals and the sequence of steps involved in each of the method to achieve the goal. For example, a document can be printed either by clicking on the printer icon or by opening the print dialogue box and selecting print options.
4. *Selection rules:* Used to predict which of the methods will be used by a particular user or in a particular situation. For example, the user may decide to use the print dialogue box when printing a specific page and use the printer icon when the document to be printed is a single page and only one copy is required.

4.2.3.3.3 Benefits of GOMS

The GOMS model is an effective method for comparative analysis of different prototypes or designs in guiding the decision between competing alternatives. For example, GOMS was used to predict the effectiveness of a proposed system in a telephone company. The results obtained showed that the proposed system would slow down telephone operators. Thus, the system was abandoned before installation, which saved the organization from implementing a potentially inefficient system [Dix et al., 2004; Preece et al., 2007].

GOMS is also useful in describing the procedures that a user must perform to operate a system, making it suitable for the design of user-manuals. GOMS can be used to generate an accurate, complete and detailed procedure for executing a specific task [Kieras, 2008].

4.2.3.3.4 Limitations of GOMS

In spite of the benefits of GOMS the application of the model is quite restricted. Some of the limitations of GOMS model are [Dix et al., 2004; Kieras, 2008; Preece et al., 2007]:

- The GOMS model can be used to predict user performance only for a small set of well-defined tasks; for example, data-entry type tasks.
- GOMS assumes that the user is an expert and does not provide for modelling user errors. This makes it difficult, if not impossible, to predict how an average or novice user will carry out tasks when using a range of systems, especially those that have been designed to support flexibility of use.
- The model assumes that a task is executed sequentially until completed. However, in reality, a typical user is often multitasking and interrupted, for example by a telephone, an incoming mail, or answering a colleague’s question.
- GOMS’ focus is on the procedural aspect of interface design but does not address other aspects of usability, such as the readability of interface text. This makes the model unsuitable for summative evaluation.
- GOMS does not account for individual differences and factors which affect user performance, such as user fatigue, mental workload, learning effects, organization and social factors.

4.2.3.3.5 The Keystroke-Level Model

The keystroke-level model is a quantitative analysis method used to predict the time an expert user will take to complete routine tasks on an interactive computer system [Card, Moran and Newell, 1980]. It aggregates a total time for the task by counting the number and duration of keystrokes, and the time spent on mental operation and system's responses. It is a simple, effective and flexible method based on the assumption that the time taken by an expert to complete a routine task depends on the time it takes to do the keystrokes [Card et al., 1980].

The keystroke-level model decomposes any give task into two phases: task acquisition and task execution. During the acquisition phase, the user develops a mental representation of the task while at the execution phase the user invokes the necessary system functionalities required to accomplish the task [Card et al., 1980; Dix et al., 2004].

4.2.3.3.6 Benefits and Limitations of the Keystroke-level Model

The keystroke-level model is a simple, yet effective method for predicting the time required to execute routine tasks. Empirical studies showed the method is capable of providing eighty percent accurate predictions [Dix et al., 2004]. However, as with GOMS model, the keystroke-level model is limited in its application area. The model also assumes the user to be an expert and does not account for user error.

4.2.3.3.7 Conclusion to Model-Based Evaluation

Model-based evaluation techniques are effective in predicting expert, error-free performance for certain computer-based tasks and can be used to compare alternative design options. They are inexpensive as user involvement is not required, and the system being evaluated may not even exist. The description of a proposed system is enough to predict the time required to complete a given task.

4.2.3.4 Interviews

4.2.3.4.1 Introduction

An interview is an evaluation method which is well-established in social science, market and HCI research [Preece et al., 2002]. They can be used in an evaluation to elicit whether an interface or a fully implemented system meet users' requirements by asking them directly. According to Cannell and Kahn [1968] (cited in Lindgaard [1994:149]), "an interview is a two-person conversation, initiated by the interviewer for the purpose of obtaining research-relevant information, and focused by him on contents specified by research objectives of systematic description, prediction, or explanation". The extent to which an interview becomes similar to a conversation depends on the structure of the interview.

Interviews need to be properly planned to gain maximum benefit. This means clearly specifying the overall aim of the evaluation. For example, is the interview being conducted as a formative evaluation during early design stage or is it a summative evaluation aimed at determining users' experience with an interface? The goal of the interview will guide the decision to select a specific interview method [Preece et al., 2007].

4.2.3.4.2 Structured Interviews

In a structured interview, the evaluator poses a series of short, clearly worded, and predetermined questions to the interviewee. This type of interview is typically used when the aim of the evaluation is clearly understood; for example, to gather feedback on users' experience after interacting with an interface. The interviewer asks all participants in the evaluation the same questions, and questions require precise answers such as yes/no responses or selection from a list. Since the evaluator usually reads aloud the question to the interviewee and records his/her response, the interview can be conducted telephonically. Structured interviews are rigid, which makes rigorous analysis of the data possible but further probing to clarify or expand an issue impossible [Lindgaard, 1994; Preece et al., 2007].

4.2.3.4.3 Unstructured Interviews

In unstructured interviews the evaluator poses open-ended questions and the interviewee can answer as fully or as briefly as s/he wants. This type of interview is exploratory in nature and can be used early in the development phase to get feedback from users on ideas for a proposed system [Lindgaard, 1994; Preece et al., 2007]. Unstructured interview can generate rich and copious data which can reveal aspects that were not previously considered by the evaluator but can be time-consuming and complex to analyze or code. In an unstructured interview, both interviewer and interviewee can direct the course of the interview and responses are influenced by the willingness and/or ability of the interviewee to provide information. Thus, there can be great variability in the type of questions asked of participants. Hence it is essential for the interviewer to prepare an outline of the main points to be covered to ensure that relevant answers are obtained for questions [Preece et al., 2007].

4.2.3.4.4 Semi-structured Interviews

A semi-structured interview combines features of both structured and unstructured interviews and is thus more focused than an unstructured interview and less rigid than a structured interview. A semi-structured interview is appropriate when the evaluator has clear goals but at the same time wish to explore any additional issues that may emerge during the interview [Lindgaard, 1994; Preece et al., 2007] . Typically, the evaluator asks a series of predetermined questions followed by open-ended type questions based on the user's response. To ensure consistencies among interviewees, the interviewer should use a guideline so that the same topics are covered with all interviewees [Lindgaard, 1994; Preece et al., 2007].

4.2.3.4.5 Focus Group Interviews

A focus group is an interview technique involving a group of three to ten representative users at the same time to discuss their needs for a proposed system or feelings about a newly deployed interface. The discussion is typically unstructured, but moderated by a facilitator who guides the procedure using a pre-planned script to ensure that relevant topics are covered. The moderator must keep the discussion on track, without hindering the free flow of conversation, encourage quiet people to participate in the discussion and prevent the verbose from dominating. A focus group interview is beneficial in that it enables sensitive and diverse issues that may not emerge using other interview methods, to be explored by participants [Preece et al., 2007].

4.2.3.4.6 Guidelines for Conducting Interviews

An interview requires careful preparation and planning. Firstly, this establishes the goal of the evaluation; the range of issues that should be covered; where the interview will be conducted; how the session will be recorded; and, how the data will be analyzed. For example, a requirement elicitation interview at early development stage is typically exploratory to provide the flexibility for further probing. Next, planning accounts for practical issues. For instance, a recorded interview session requires transcribing before analysis and this can be time-consuming and the data may be difficult to code or categorize [Lindgaard, 1994; Preece et al., 2007]. Interviews should also be preceded by a pilot study to uncover any problem and gain the requisite experience for conducting an interview [Preece et al., 2002].

The questions that are posed and how they are phrased in an interview should be guided by the length, clarity, and relevance of the question [Lindgaard, 1994]. Preece et al. [2007] proposes the following guidelines:

- Long questions should be avoided as they are difficult to remember.
- Compound questions can be confusing; such questions should be decomposed into two separate questions that can be answered separately.
- Technical terms should not be used as they may not be understood by the interviewee.
- Leading questions, which presuppose a particular response from participants, should be avoided.
- Interviewers should be aware of their own personal biases, making sure such biases are not reflected in questions posed.

Typically, an interview session consists of the following stages [Preece et al., 2007]:

- An introductory session, in which the interviewer introduces him/herself; explains the purpose of the interview; reassures the interviewee; explains how data collected would be used; seeks the interviewee’s permission for the use of the data, including recording the session, if applicable; and, signature on an informed consent. The outline of this session should be consistent for all participants.
- The warm-up session, in which the interviewer asks simple, general questions, for example, demographic questions like age and gender.
- The main session, in which the interviewer asks questions that address the main focus of the interview and are posed in ways that depend on the response of each participant.
- A cool-off period, in which the interviewer asks further simple questions, to ease any tension that might have arisen and gives interviewees the opportunity to provide any other information they may wish to add.

- The closing session, where the interviewer thanks the interviewee for participating in the evaluation and switches off the recorder to signal the end of the interview.

4.2.3.4.7 Conclusion to Interviews

An interview is a good method for eliciting direct information about users' experience with an interactive system. However, participants' responses are subjective, especially when using unstructured interviews. In addition, participants may provide answers they think the interviewer wants to hear. While it may be impossible to avoid participant subjectivity, it is essential to be aware of them.

4.2.3.5 Questionnaires

4.2.3.5.1 Introduction

Questionnaires, or surveys, [Ozok, 2008; Shneiderman, 1998], are another method for involving users in the evaluation process and one of the primary methods of data collection in HCI field due to the relative ease of administration and the potential to reach a wider target audience [Dix et al., 2004; Ozok, 2008; Preece et al., 2007]. Questionnaires involve a set of questions administered via a computer or paper-and-pencil environment [Ozok, 2008] and are similar to interviews in terms of the closed or open nature of the questions. However, they are not as flexible as interviews because questions are fixed and further probing is impossible.

Questionnaires can be used as a stand-alone evaluation method or as part of another evaluation method, such as following usability testing to measure the subjective views of participants on the application just tested [Dumas, 2003; Preece et al., 2007].

It is more difficult to develop questions for questionnaires than for a structured interview. Respondents may sometimes interpret a question differently which means they provide answers which are incorrect from the perspective of the evaluator [Nielsen, 1993]. Typically, the evaluator is not available to clarify ambiguous or unclear questions [Preece et al., 2007], this may compromise the quality of the data. Thus, like interviews, successfully administering questionnaires depends on proper advance planning. This establishes the purpose of the evaluation and the method of analyzing the data generated [Dix et al., 2004].

4.2.3.5.2 The use of Questionnaires in HCI Research

In HCI questionnaires are used to collect information regarding users' attitudes, preferences and product evaluation. Although earlier approaches involved the use of paper-and-pencil, in recent years, questionnaires are often designed to be completed online [Ozok, 2008].

According to Ozok [2008], questionnaires used in HCI research can be classified into three main categories:

1. *Questionnaires for user evaluation* are used to collect information regarding the degree to which a system, product or application meets users' needs, goals or expectations for summative evaluation purposes. Questions are asked that seek to establish the overall impression of users about a

specific product or application, what types of problems users encounter while using it, and the ease of learning and using the application.

2. *Questionnaires for user opinion* gather information from users and potential users about the requirements for upgrades to an existing system or a new one. These types of questionnaires fit those used for formative evaluation as part of user-centred approach to design and development of interactive systems.
3. The third category of questionnaires constitutes those that do not fall within the first two categories. Examples of these are those used to collect demographic information such as the age, sex, education background of respondents, and the frequency of use of a particular application.

Although Ozok [2008] have categorized questionnaire types into three main groups, any given questionnaire can include questions from a combination of the three categories.

4.2.3.5.3 The Design of Questionnaires

The design of a valid and reliable questionnaire for usability evaluation requires a substantial amount of effort and specialized skills. Since the expertise of many usability experts lies elsewhere Dumas [2003] recommends the following steps to create an effective questionnaire:

- Creating an initial extensive set of questions that focus on the specific attitudes or opinions that are to be measured.
- Eliminating poor questions in the set by first asking a number of potential users to complete the questionnaire and then calculating the correlation and variances between each question and the correlation of the total scores of all questions. Questions with low correlation and those with small variances should be removed. Furthermore, if any two questions have high correlations, one of them should be removed as it indicates that both are measuring the same thing.
- Evaluating the reliability of the questionnaire in measuring a given quantifiable phenomenon consistently [Ozok, 2008], by *test-retest*. That is, the same group of respondents are asked to complete the questionnaire twice, but with a time lapse between these processes, so that they are unlikely to remember their previous responses.

  Ozok [2008] identified two measures of reliability in HCI research, *internal* and *inter-rater* reliability. Internal reliability measures the degree to which participants understand the questions while inter-rater reliability measures the consistency among participants' responses to the same question. Clearly, internal and inter-rater reliability cannot be measured for evaluation studies that include open-ended or semi-structured questions. Such questionnaires may also seek participants' opinion, which may not necessarily be the same across respondents [Ozok, 2008].
- Evaluating the validity of the questionnaire in measuring what it is supposed to measure [Ozok, 2008]. Questionnaire validity is the most vital aspect of a questionnaire but is the most difficult to assess. One way is to determine the correlation between both typical users and experts in testing the application under evaluation. A high correlation between the scores of the experts and the users suggests the questionnaire is valid [Dumas, 2003].

4.2.3.5.4 Types of Questions in Questionnaires

Dix et al. [2004] identifies five types of questions that can be included in a questionnaire.

1. *General questions* which collect background information about the respondent. They include questions relating to demographic information like age, gender and occupation, and previous experience of the respondent. General questions can be asked using open-ended, multi-choice or scalar questions (see below). These types of questions falls within the third category of questionnaires identified by Ozok [2008].
2. *Open-ended questions* are similar to those asked in an unstructured interview where the respondent is free to provide a subjective view on the product being evaluated. Open-ended questions are difficult to analyze rigorously, but are useful in identifying aspects that were not previously considered by the evaluator.
3. *Scalar questions* allow respondents to provide judgement about a statement. Numeric scales are used to indicate an agreement or disagreement with the statement. The granularity of the scale in general varies. For instance, a scale of one to three is very coarse, and a scale of one to ten is finer. Scales of one to five or one to seven provide sufficient levels of differentiation to the respondents. Odd-numbered scales allow respondents to stay ‘neutral’ and even-numbered scales are used if ‘neutral’ will not focus the study. Scalar questions often make use of Likert scales [Dix et al., 2004; Preece et al., 2007].
4. *Multi-choice questions* are similar to those asked in structured interviews and typically collect information on participants’ previous experience. They enable respondents to provide specific answers from a list of possible options where they can select just one option or as many as are applicable.
5. *Ranked questions* order the items in a list and are useful for indicating participant’s preferences.

In general, the rigour available in analyzing scalar, multi-choice and ranked questions makes them preferable. Because they provide respondents with a variety of alternatives to select from, they may help respondents to complete the questionnaire and thus facilitate a higher return rate. Some respondents may find it difficult to answer open-ended questions, which decrease the probability of returning the questionnaire.

4.2.3.5.5 Off-the-shelf Questionnaires

Many usability experts do not have the time and skills required to develop valid and reliable questionnaires and purchase professional off-the-shelf alternatives. Specialists develop and test the validity and reliability of off-the-shelf questionnaires such as:

- *Software usability scale (SUS):* The SUS was developed at Digital Equipment Co Ltd in the United Kingdom “as part of the usability engineering programme in integrated office systems development” [Brooke, 1996]. It is aimed at addressing the need for an evaluation method that is simple enough to be quickly administered, yet reliable enough in the subjective attributes being measured. SUS can be used as a stand-alone evaluation method or as part of another evaluation method, and can be used for any product, not just software. It is administered after the user has interacted with the application but prior to any form of discussion about the product. SUS consists of ten Likert scale questions with five point rating, and the total score can range between zero and hundred [Brooke, 1996].

- *Computer user satisfaction inventory (CUSI):* CUSI is a 22-question questionnaire developed to measure user attitude to software applications. The 22 questions are grouped into two subscales, the first measuring how well the respondent likes the software and the second measuring how competent s/he is in using the application to complete a specific task [Dumas, 2003].
- *Questionnaire for user interaction satisfaction (QUIS):* QUIS was developed for use as a stand-alone evaluation method to provide subjective measure of several aspects of interactive interfaces. It consists of two groups of questions: the general questions, which measure the entire application; and the detailed questions, which measure the application's interface. QUIS is available in two forms: the short form, consisting of 26 questions; and the long form, which consists of 71 questions. In practise, evaluators often use a subset of the questions. Each question has a nine point rating scale. QUIS is one of the most widely used questionnaires for evaluating interactive interfaces [Dumas, 2003; Preece et al., 2007; Shneiderman, 1998].
- *Software usability measurement inventory (SUMI):* SUMI was developed by the Human Factors Research Group at the University College of Cork for use as a stand-alone method to evaluate software applications with full functionality and those still under construction. SUMI is different from the questionnaires discussed above, as it uses a set of fifty statements rather than questions. Users respond to these with *agree*, *undecided*, or *disagree*. The statements are grouped into six subscales, namely, global, efficiency, affect, helpfulness, control, and learnability. SUMI is typically administered following several usage of the application being evaluated by the user and it can be completed within five minutes. SUMI has been tried and tested and it is recommended that a qualified psychometrician do the test scoring [Dumas, 2003].
- *Measuring the usability of multi-media systems (MUMMS):* MUMMS was also developed by the Human Factors Research Group at the University College of Cork for evaluating the usability of multi-media applications. The questionnaire has the following sub-scales:
    - How much the application captures the user's emotional responses?
    - The extent to which the user feels s/he is in control and not the application.
    - The extent to which the user can use the application to accomplish his/her goals.
    - The extent to which the application seems to assist the user.
    - The ease with which the user can learn to use the application.

  The second version of MUMMS is being developed, and the developers are planning to include another subscale called *excitement*, which measures the extent to which the user feels 'drawn into' the world of the application [Human Factors Research Group, n.d; Preece et al., 2002].

4.2.3.5.6 Conclusion to Questionnaires

Questionnaire is an effective method for evaluating usability. A large number of respondents can be reached in less time compared to interviews; data generated can also be analyzed more rigorously but the distribution and return rates can be problematic. Design of valid and reliable questionnaires requires specialized skills and a considerable amount of time. Off-the-shelf questionnaires that have been tried and tested offer solutions to such requirements.

One major shortcoming of questionnaires is that respondents sometimes provide answers they deem to be socially acceptable, and this may not adequately reflect the usability of the particular application [Nielsen, 1993].

4.2.3.6 Observations

4.2.3.6.1 Introduction

Observation is an evaluation method where real users are observed while interacting with the target application. The method can be used early in the development process as part of task analysis or to evaluate the usability of a fully functional application following deployment in a summative evaluation. Observing real users interact with an application can reveal unexpected ways of use, the result of which can feed into subsequent versions of the application [Dix et al., 2004; Lindgaard, 1994; Nielsen, 1993; Preece et al., 2007].

Observation on its own may not be sufficient to determine the usability of an application as it may be difficult to determine the reasoning behind users' actions. For this reason users are often asked to talk through their actions by 'thinking aloud' [Dix et al., 2004]. The think-aloud method is discussed in section 4.2.3.7.7.

4.2.3.6.2 Types of Observation

Observation can be done by a researcher in a controlled environment as a component of another evaluation method or in a natural setting [Dix et al., 2004; Dumas, 2003; Lindgaard, 1994; Preece et al., 2007]. In each of these settings, users can be observed directly or indirectly. Each type of observation is discussed below [Preece et al., 2007]:

- *Direct observation in the field* the researcher, who either assumes the role of an insider (participant observer) or an outsider (passive observer), observes the user while carrying out normal or routine activities either at home or at the 'workplace'. Direct observation in the field can reveal details that are difficult to obtain using other evaluation methods, as it enables the researcher to see the context of use of the application. However, direct field observation can be difficult due to high levels of noise or constant interruptions from colleagues. Furthermore, the method can generate unstructured and sometimes irrelevant data that is difficult to analyze. These factors make it imperative that proper planning is conducted beforehand.

  Field observation data is typically recorded using hand-written notes, audio and video recordings. Analysis of observational data should be done at the end of each day; notes and recordings should be checked and a summary of the day's events written up so that valuable information is not lost or forgotten.

Field observation is obtrusive and may constitute the invasion of participants' private space, thus informed consent of people being observed should be sought obtained.

- *Direct observation in a usability laboratory* the evaluator observes participants as they carry out specific tasks while using the application being evaluated as part of a usability testing process. The observation can occur via a one-way mirror or on a computer monitor or television screen. The observational data is stored for later review to get insight into users' emotive reactions. Usability testing is discussed in section 4.2.3.7.
- *Indirect observation* involves the recording of users' activities for later review. This method can be used when direct observation can interfere with users' activities or when it is impossible for the observer to be present at the location of observation. Using diaries or journals, users make notes important points, for example, time spent on tasks and components that were easy or difficult to use as they go about their activities. This approach to indirect observation is advantageous, since the technique is inexpensive and requires no special equipment or expertise. The method is not without disadvantages, however. Participants may omit to record important information or even exaggerate the occurrence of certain events. To be effective, an easy-to-complete template can be used and some form of incentives may be provided.

  Another form of indirect observation is interaction logging. Using this approach, a software application captures the user's activities, such as key presses and time spent using the system, in a log for later review. This can be done as part of usability testing or to monitor the pattern of use of a system. Interaction logging is unobtrusive but can constitute the invasion of privacy, since users are often unaware of being monitored.

4.2.3.6.3 Planning a Field Observation

Adequate planning is imperative prior to implementing any evaluation technique. It is even more so with observation in the field where events can be dynamic. There should be a well articulated goal for the observation to ensure that important events are not missed.

A simple framework that is based on the 'who' (the person using the target application), 'where' (the environment of use) and 'what' (the task that is being carried out using the application), can be effective in focusing the evaluation process [Preece et al., 2007].

The decision should be made upfront regarding the degree of participation that will be adopted, that is, whether the evaluator will be an active or passive observer. The goal of the evaluation and other ethical considerations will influence this decision.

Other decisions include how data will be recorded, for example, whether or not video cameras will be used, how to gain acceptance by study participants, and the manner in which issues of sensitivity will be handled [Preece et al., 2007].

4.2.3.6.4 Benefits and Limitations of Observation

Observations, especially direct field observation, enable evaluators to see the context of use of the particular application being assessed. Evaluation with real users can also reveal problems that may impact on user tasks, some of which might be overlooked by other evaluation methods involving expert analysis.

Observation as a stand-alone evaluation method may be inadequate to assess the usability of an application. This is because it is difficult to determine the reason why users make certain decisions. Some of the limitations to the use of observation as a sole evaluation method include the following [Dumas, 2003]:

- Because the observer has limited control over the flow of events as they unfold, it becomes difficult to determine the cause of a particular behaviour.
- It may be difficult to record the true behaviour of participants, as people often change the way they behave when they become aware of being observed (the so called 'Hawthorne effect').
- It is possible for the observer to see what s/he wants to see, which may affect the validity of the observational data.

4.2.3.6.5 Conclusion to Observation

Observation is a useful method for getting users' feedback on the usability of an application but the method has a number of limitations. It is more beneficial to combine observation with another evaluation method such as usability testing or heuristic evaluation.

4.2.3.7 Usability Testing

4.2.3.7.1 Introduction

Usability testing is an approach to user-centred design that places emphasis on real users, the tasks they want to accomplish and the context of use of the application. Rubin [1994:25] describes usability testing as "a process that employs participants who are representative of the target population to evaluate the degree to which a product meets specific usability criteria". The goal of usability testing is to determine the extent to which the target application meets specific measures of usability and not the user's ability [Preece et al., 2007].

Usability testing started in the early 1980s with the growth of personal computers and applications, changing the usage profile from enthusiasts and scientists to the casual users [Dumas and Fox, 2008; Shneiderman, 1998]. During this early period, usability testing was expensive and time-consuming as it required usability specialists with a variety of skills in human cognition, experimental psychology, human factors experience, and as many as thirty to fifty participants to conduct tests [Barnum, 2002]. The high costs, intensive time, and specialized skills required to conduct usability testing did not provide enough incentives for organizations to conduct usability testing.

Nielsen [1994a], proposed a less expensive and relatively effective approach, termed *discount usability testing.* This approach does not require sophisticated equipment typical of a controlled usability laboratory. The approach requires between three to five participants; one of the observers uses a stopwatch to record time spent on tasks and record essential findings on a form. Other observers record user actions, with comments for later discussion [Barnum, 2002]. This makes it feasible to incorporate usability testing early into the development process without time and cost overruns, thus shifting the focus to the user.

Usability testing typically has six features; any usability testing lacking even one of these key characteristics cannot be regarded as usability testing [Dumas and Fox, 2008]:

1. The focus of evaluation is on usability.
2. The participants in the evaluation are end users or potential end users.
3. There is a specific artefact to evaluate, for example a product design, a prototype or a fully functional system.
4. The participants think aloud as they perform real tasks.
5. The data is recorded for subsequent analysis.
6. The results of the test are communicated to relevant stakeholders.

Based on the position of Dumas and Fox [2008], one cannot for example, regard as usability testing an evaluation where co-workers are used as evaluation participants if they do not fit the user profile for the application, nor can one label as usability testing an evaluation which has as its main goal the prediction of the likelihood of participants purchasing the product.

4.2.3.7.2 Planning the Usability Testing

Proper planning should precede usability testing as it forms the foundation for the entire testing process. The purpose of the test should be well articulated right from the beginning, since this will guide the conduct of the test. A test plan should generally address the following issues [Nielsen, 1993; Rubin, 1994]:

- What is the purpose of the evaluation?
- Where and when will the test take place?
- Who will be the test participants? How many of them are required and how will they be recruited?
- What equipment/software would be required?
- What usability metrics will be taken and how will they be analyzed?
- Who will be the test administrator?
- How will results from the test be communicated to relevant stakeholders?

Clarifying the purpose of the test will impact on the type of testing to be performed. Testing could be done as part of the formative evaluation process, for example, to guide the design and improvement of the user interface. Usability testing as a summative evaluation method is done to collect both measurable and subjective data for a fully functional application.

Usability testing should be preceded by a pilot, allowing the evaluator to put all the planning activities into practice and to try out the methods and the set of intended tasks. Conducting a pilot test prior to the main study also provides the opportunity to ensure that the equipment is functioning satisfactorily [Barnum, 2002].

4.2.3.7.3 Recruiting Test Participants

Participants in usability testing should be representative of the intended users of the application. To find potential candidates for a test, a user profile is established, including age group, computer usage and educational background [Barnum, 2002; Rubin, 1994].

Participants can be recruited directly or through an agency. When recruiting participants directly, appropriate means of searching include the company database (if there are pre-qualified users), product specification document (which usually includes the intended users of the product), customer lists obtained from sales and marketing people, college/university campuses, and qualified friends/relatives. To ensure that potential participants meet the user profile, a screening questionnaire is developed for them to complete. It may also be necessary to offer some incentives in the form of token gifts. Finally, test participants should be contacted, preferably by phone, a day or two before the test to confirm the appointments [Barnum, 2002; Rubin, 1994].

4.2.3.7.4 Ethical Considerations

Participants in every evaluation should be treated with respect and dignity. Although usability testing might not expose participants to physical dangers, the controlled nature of the test environment, the presence of recording cameras and participants' awareness of being watched, can be a source of distress to some participants.

Participants may sometimes feel pressured to perform well, even when they have been told that it is the application that is being tested, not them. This applies to both novice and advanced participants. The test administrator should make the participants feel as comfortable as possible during and after the evaluation Terms like 'subjects' and 'guinea pigs' should never be used to refer to participants.

Participants should be given an informed consent document to read and sign. The form should specify the purpose of the test, explain what will happen during the test, and mention the presence of recording cameras in the usability laboratory. The document should explain how data collected would be used.

Participants should be informed of the voluntariness of their participation, the right to ask questions or withdraw from the test at any time without any negative consequences [Dumas and Fox, 2008; Nielsen, 1993; Preece et al., 2007; Shneiderman, 1998].

During the testing session, the evaluator should avoid interfering and allow the participant to discover solutions to problems by him/herself to prevent biasing the result. However, when it is obvious that a participant is struggling with the task, the evaluator should intervene by providing hints on how to continue.

4.2.3.7.5 Usability Test Metrics

Usability test metrics generally include two types of data; quantitative data, aimed at measuring user performance, and qualitative data, a subjective measure of the affective behaviour of users. The purpose of the evaluation should guide the type of data collected. Common quantitative performance metrics include the following [Barnum, 2002; Nielsen, 1993]:

- Time taken to complete a task.
- Number of tasks completed, without help and after help.
- Number of tasks not completed.
- Number of errors, whether recoverable or not.
- Number of participants making a particular error.
- Time taken to recover from an error.
- Number of features not used.
- Time spent navigating.
- Number of calls for assistance.

In addition to the usability metrics listed above, qualitative measurements are usually taken. These includes participants' verbal expressions of frustration or satisfaction, observation of frustration or satisfaction by the test administrator, and participants' opinion about the application [Barnum, 2002; Dumas and Fox, 2008; Nielsen, 1993; Preece et al., 2007]. A participant's subjective opinion is typically assessed through post-test questionnaires or interviews.

4.2.3.7.6 Supplementary Techniques to Usability Testing

Some techniques that can be used to support or supplement usability testing include:

- *Think-aloud:* A data gathering method used in usability testing and observation where participants are encouraged to verbalize their thoughts, feelings, expectations and decisions as they interact with the application being evaluated. This assists evaluators to gain insight into the reasoning behind users' actions [Barnum, 2002; Dix et al., 2004; Preece et al., 2007].

  A participant can think aloud while carrying out the specified tasks *concurrent think-aloud* or after the completion of the test session *retrospective think-aloud* where the participant carries out the tasks in silence. Thereafter s/he is shown the recordings and asked to talk through the reasoning behind his/her decisions during the task [Dumas, 2003; Lindgaard, 1994].

  Many people find thinking aloud to be unnatural. In a study by Adebesin, De Villiers & Ssemugabi [2009] to analyze time usage patterns in the evaluation of an e-learning tutorial, single participants in the study struggled to think aloud despite prior coaching on the use of the technique. A number of authors, including Kahler, Kensing and Muller [2000], Nielsen [1993], Wildman [1995] and Wilson [1998] have acknowledged the think aloud problem in single participant usability testing.

To address the unnaturalness involved in thinking aloud, a variation of the method, called co-discovery or co-participant testing, can be used.

- *Co-discovery:* Two users collaborate with each other while exploring the application being evaluated. The idea is that they verbalize their thoughts as they interact with each other and the application, using a single workstation. It is a variation of the think-aloud method, but instead of a single user expressing his/her thoughts and feelings while interacting with an application, the verbalizing in this case is more natural, because it involves a conversation between two people [Nielsen, 1993; Wildman, 1995].

  Selecting and paring participants for co-participant testing can be difficult since the level of expertise of the participants has to be considered. Nielsen [1993] recommends the same level of expertise for both participants, but Kahler et al. [2000] advocate pairing participants with different levels of expertise. In general, it is better to pair participants who already know each other, for example, friends, family members, or co-workers [Kahler et al., 2000], provided they meet the user profile for the target application.

  When compared to the classical think-aloud, co-participant testing has a number of benefits, including the following [Wilson, 1998]:

  - Co-discovery is useful at early design stages for capturing the conceptual model and compare alternative designs.
  - It is more natural for participants to think aloud as they collaborate, thus producing more comments.
  - Testing is faster, because it requires half as many test sessions to reveal the same number of problems.
  - Co-participant testing reduces the need for the test administrator to be present in the evaluation room with the participants, thus limiting the probability of him/her biasing the participants [Wildman, 1995].
  - With co-participant testing, the number of calls for help is reduced and the test administrator rarely need to intercede with help because both participants assist each other in resolving problems [Adebesin et al., 2009; Wildman, 1995].

  Co-discover has a number of drawbacks, including [Wilson, 1998]:

  - Differences between participants' level of expertise, learning styles, verbal, and cultural background can affect feedback from participants. For example, if one of the participants is verbose and the other more reserved, then the verbose individual will dominate the session.
  - Because differences in level of expertise may affect the result, the screening of participants is more stringent than in single participant testing.
  - Co-participant testing can generate a huge amount of qualitative data. This makes the analysis and interpretation more complex than single participant testing.

  - Co-participant testing requires the use of twice as many participants as single participant testing. Its use should thus be limited to evaluations where it is easy to recruit a greater number of participants [Nielsen, 1993].

- *Cooperative evaluation:* This is a variation of the think-aloud technique where the evaluator interacts with the participant and encourages him/her to ask questions or clarifications whenever a problem is encountered. The evaluator can also ask the participant questions on why certain decisions are made. This approach is less intimidating especially to the novice participant, who is assured of 'help-in-hand' whenever the need arises [Dix et al., 2004].

- *Eye tracking:* A technique for monitoring the movement of participants' eyes on different screen regions during usability testing. Eye tracking technology enables evaluators to isolate the exact source of usability problems [Dumas and Fox, 2008; Pretorius, Calitz and Van Greunen, 2005]. Using the technique, it is possible to determine whether screen instructions or important information is perceivable and comprehensible.

  Traditional eye tracking technologies were awkward and invasive, requiring physical contact with the participant. For example, some eye-trackers require contact lenses to be inserted into the participant's eyes and others involved the monitoring equipment mounted on the participant's head. In recent years, the technology has advanced considerably. Modern eye-trackers, for example, the newer versions of the *Tobii* eye tracker, are not obtrusive. This makes it easier for participants to focus on the test session.

  Incorporating eye tracking into usability testing can provide designers with information that would be difficult to obtain using only the traditional measures of usability. For example, a simple usability testing metric showing increased task duration might indicate that the participant had problems with the interface, but it can be difficult to determine the exact component on the interface that caused the problem. When the testing session includes eye tracking, a longer fixation on a particular component might indicate that the participant had problem understanding the use or meaning of that component.

  Eye tracking also has a number of drawbacks, such as calibrating the equipment with the participant's eyes before testing sessions; incessant head movements might result in inaccuracies. In addition, the technique cannot be used in participants wearing bifocal lenses [Dumas and Fox, 2008].

4.2.3.7.7 Conclusion to Usability Testing

Usability testing involves measuring the performance of typical users while carrying out specified real-world tasks on the target application. Formal usability testing is typically carried out in a specialized usability laboratory with monitoring and analysis equipment and software. Proper planning should be done before conducting usability testing, and preceded by a pilot test. Maximum benefit can be obtained from usability testing when combined with techniques such as think-aloud and eye tracking.

4.2.4 Factors Influencing the Selection of an Evaluation Method

Sections 4.2.3.1 to 4.2.3.7 examined seven usability evaluation methods. Each of these methods requires resource commitment in the form of time, money, equipment and expertise. The stage at which evaluation is done will also influence which method is selected. In practise, evaluation methods are combined to triangulate data and get multiple views on the usability of the application [Preece et al., 2007]. While there are no strict rules for selecting an appropriate evaluation method, certain factors should be considered and used as guidelines for selecting evaluation method(s) that suit a particular situation. The eight factors that should be considered are [Dix et al., 2004]:

1. *The stage at which evaluation is carried out:* Evaluation can be formative or summative (section 4.2.2). Evaluation should be incorporated into the development lifecycle; the result of evaluation should guide further design and development. During early development stage, it should be easy to implement a formative evaluation method quickly without consuming too many resources. In the later stage of implementation, summative evaluation has to be more comprehensive and may require user involvement. Evaluation method like the heuristic evaluation method (section 4.2.3.1), is suitable for formative and summative evaluation.
2. *The style of evaluation:* Evaluation can be conducted in a controlled environment or the natural setting of application usage. The development stage at which evaluation is done will influence the style of evaluation. Evaluation in a laboratory makes it possible for the evaluator to control the evaluation process but the natural context of use is lost. Evaluation in the natural environment of use can reveal details that are difficult to obtain in a formal laboratory setting, as it enables the researcher to see the context of use.
3. *Level of subjectivity or objectivity of the technique:* Evaluation methods, such as heuristic evaluation and cognitive walkthrough (section 4.2.3.2) are highly subjective and rely on the expertise of the evaluator. Unstructured interviews (section 4.2.3.4.3) and think-aloud (section 4.2.3.7.6) can generate rich data if used properly. Other methods like usability testing (section 4.2.3.7) in a usability laboratory can produce repeatable results if used correctly with homogenous subjects.
4. *Type of data generated:* Evaluation techniques generally produce quantitative or qualitative data. Methods like questionnaires (section 4.2.3.5) provides quantitative data that can be rigorously analyzed, while methods like unstructured interviews and think-aloud results in qualitative data, which is more difficult to analyze. The level of objectivity or subjectivity of the method also influences the type of data generated.
5. *The information provided:* The level of information provided can range from low-level to high-level information. This is influenced by the goal of the evaluation. Sometimes evaluation aims for low-level information, for example, the most appropriate font size for a given interface. Others seek high-level information, for example, the extent to which an implemented interface meets users' requirements. Methods like usability testing in a controlled environment are ideal for low-

level information while others, like interviews and questionnaires, are useful for obtaining high-level information.

6. *Immediacy of response:* The urgency with which the result of evaluation is required will influence the decision to choose a particular evaluation method. Direct observation (section 4.2.3.6) and think-aloud can produce immediate results because the user's actions and verbalizations are recorded as they occur. This can provide valuable information instantly. Other methods, like indirect observation using diaries, are dependent on the user remembering to record important events as they occur for the purpose of subsequent analysis by the evaluator.
7. *Level of intrusiveness:* The level of intrusiveness of a particular technique is closely related to when a result is produced. Methods that provide immediate response tend to be intrusive, which may lead to users changing their behaviour.
8. *Resources required:* The availability of the required resources for a given method will influence its use. Techniques like usability testing require sophisticated equipment and user involvement. Choosing between two competing methods should be guided by the expertise of the evaluator and the effectiveness, efficiency, and ease of use of the method.

4.2.5 A Framework for Evaluating Interactive Systems

In section 2.2.2.1, I made the case for evaluation. Evaluation should not be seen as a separate phase of the development lifecycle, but an integral part of development which should occur throughout. When problems are identified early, the costs and efforts to correct them are far less than at later stage.

As discussed in sections 4.2.3.1 to 4.2.3.7, successful evaluation depends on proper planning; this allows the evaluator to anticipate potential problems before they arise so that contingency plans can be put in place. The selection of specific evaluation method(s) will depend on the product being evaluated, the goal of the evaluation and the ability to implement the chosen method(s). Practical constraints that could impact on the evaluation should be considered. Tight deadlines, low budgets and inability to recruit suitable participants will affect the conduct of evaluation.

Evaluation frameworks, such as the DECIDE framework recommended by Preece et al. [2007], can provide practical guidelines for the conduct of usability evaluation. The DECIDE framework make use of a checklist of items that assist in planning the evaluation and the essential issues that should be considered. The framework consists of the following items:

1. Determining the *goals*.
2. Exploring the *questions*.
3. Choosing the *evaluation approach and methods*.
4. Identifying *practical issues*.
5. Deciding how to deal with the *ethical issues*.
6. Evaluating, analyzing, interpreting, and presenting the *data*.

The framework should be adapted to suit the prevailing circumstances of the evaluator. Sections 4.2.5.1 to 4.2.5.6 briefly describe each of the items in the framework.

4.2.5.1 Determining the Goals

Every evaluation should be goal-driven. Clarification of the goals is the first step in planning an evaluation. An evaluation to ascertain whether users' requirements have been adequately captured will have different goals to one that seeks to determine which of two design alternatives to select. Identifying the goal will guide the selection of appropriate evaluation method(s). For example, usability testing might be appropriate for an evaluation that aims to measure the degree to which specific usability goals have been met.

4.2.5.2 Exploring the Questions

Merely stating the goals for an evaluation is not enough to make them operational. Questions relevant to the achievement of the goals must be answered. Some of these questions might be high-level, for example: To what extent does an interface meet users' requirement? Others might be more specific, for example: What font size is the most appropriate for a given interface? Each question should be broken down into specific sub-questions at a level that is sufficient to focus the evaluation.

4.2.5.3 Choosing the Evaluation Approach and Methods

After identifying the goals and the questions that should be answered by the evaluation, the next step is to select an appropriate evaluation approach. As discussed in section 4.2.4, the decision to choose any particular method(s) will be influenced by a number of factors. Some of these factors include the stage at which evaluation is done, whether or not an evaluation should be carried out in a controlled environment, and resource availability. Trade-offs and compromises may also have to be made. For example, the method that seems to be the most appropriate may not be practically implementable due to time or budget constraints.

4.2.5.4 Identifying the Practical Issues

There are a number of practical issues that should be considered before conducting evaluations. Issues that should be considered include access to appropriate users, facilities and equipment, the practicality of evaluation considering time and budget constraints, and evaluators' expertise.

User-based evaluation requires availability of users who are representative of the target population. Users' levels of expertise, age, cultural diversity, educational experience, and personal differences have to be taken into account, depending on the type of application being evaluated. Another aspect that needs careful consideration is how the users will be involved. Tasks given to users in a usability testing evaluation should be representative of those for which the application will be used in real life.

When an evaluation requires the use of equipment, several practical issues need to be considered. For example, when using video to record an interview, the evaluator should consider how the recording will be done, the number of cameras required, and where cameras will be positioned for effective recording without making the participant uncomfortable. Where equipments are shared by several evaluators from a central pool, then the necessary equipment have to be booked in advance to ensure availability when required.

Another issue that need careful consideration is schedule and budget constraints. While it might seem ideal to test twenty users, for example, this may not be feasible if the available resources will not allow such extensive evaluation. Participants' evaluation sessions should be planned ahead although unforeseen circumstances may result in change of plan.

Every evaluation method requires some level of expertise. For example, conducting formal usability testing in a laboratory requires expertise in the use of specialized recording equipment. Quantitative studies, where statistical analysis will have to be performed, require the evaluator to possess the skills or engage the services of a statistician, and expert evaluation requires the availability of appropriate experts.

4.2.5.5 Deciding How to Deal with Ethical Issues

As discussed in section 4.2.3.7.4, evaluators should adhere to ethical practise and code of conduct when studies involve human participants. Participants should be treated with dignity and respect. Adequate information should be provided regarding the purpose of the evaluation. Personal information which could reveal the identity of participants should be kept confidential.

4.2.5.6 Evaluating, Interpreting, and Presenting the Data

Finally, a decision should be made about what data to collect to answer the study questions, how the data will be analyzed, and how the findings will be presented. The chosen evaluation method(s) will influence the type of data collected. Other issues that need to be considered include whether or not to treat data statistically, the reliability and validity of the selected method, and whether or not there are biases that could potentially distort results.

4.2.6 Usability Evaluation Methods Applicable to the Digital Doorway Context

Having examined the various methods for evaluating the usability of interactive systems, the next appropriate question is which of these methods can be used in a summative evaluation of applications installed on the DD. The heuristic evaluation method, discussed in section 4.2.3.1 is a suitable method for summative and formative evaluation, provided that appropriate evaluation heuristics are used. Its ability to uncover large numbers of potential usability problems and the relative ease with which I could recruit expert evaluators makes the method appropriate for this study.

Cognitive walkthrough (section 4.2.3.2) is also a method that can be used in a summative evaluation. However, the method's assumption that evaluators possess skills in cognitive theory and its focus on the evaluation of only the learnability aspect of an application makes it inadequate for evaluating the DD.

The two model-based evaluation methods (GOMS and the keystroke-level models) are used to predict the performance of expert users. DDs are aimed at users with little or no computer experience. This makes these methods inappropriate for the DD environment.

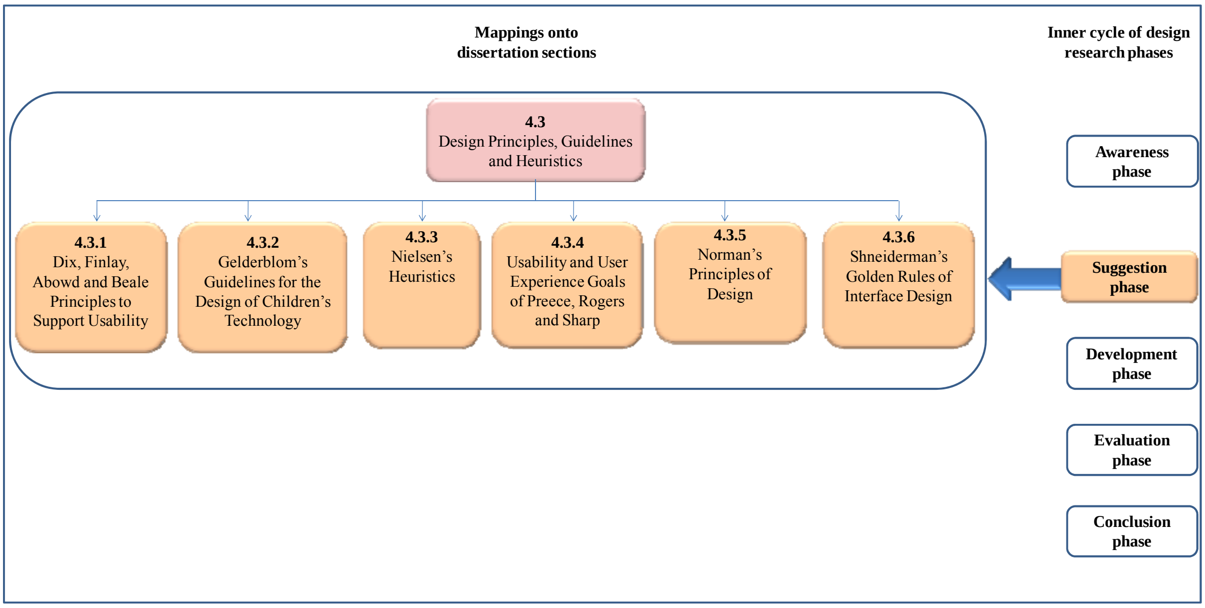
Questionnaires and interviews are evaluation methods requiring the involvement of end users. Both methods can be used in formative and summative evaluation. The ability to evaluate the usability of interactive systems from users' perspectives makes these methods appropriate for evaluating the DD. While acknowledging that any of the two query evaluation methods could be utilized in evaluating the DD, I have selected the questionnaire method because of the need to minimize disruptions to learning activities as much as possible. Conducting interviews after each evaluation session at the local school where the field evaluation was conducted could affect the number of evaluation sessions per day, since the sessions took place after school hours. This could result in the evaluation extending for a longer period.

User observation in a natural environment is advantageous as the context of use is retained. DDs are deployed into user communities around South Africa; this method is capable of revealing user problems that may be difficult to obtain using other evaluation methods like the heuristic evaluation method.

Usability testing in a controlled environment is also a method for detecting usability problems from users' perspectives. The method is highly expensive and requires the availability of sophisticated equipment. Although as a registered UNISA student I have access to a well-equipped usability testing laboratory, practical and logistical constraints, discussed in section 3.4.1.1, made usability testing inappropriate for this study.

4.3 DESIGN PRINCIPLES, GUIDELINES AND HEURISTICS

Mapping of section 4.3 to the design research phases



In section 2.2.2.2, I introduced design principles as a tool that could be used to evaluate the usability of interactive systems. This section maps onto the suggestion phase of the inner cycle of the research design process where existing usability design principles, guidelines and heuristics are reviewed to determine their appropriateness for the formulation of application-specific heuristics for the DD. This is to achieve one of the objectives of this study: To develop an instrument that can be used to evaluate the usability and direct-accessibility support provided in the DD. The guidelines that are found to be applicable will be highlighted in the relevant sections where they are discussed.

4.3.1 Dix, Finlay, Abowd and Beale Principles to Support Usability

Dix et al. [2004] provide a list of comprehensive, but non-exhaustive, design principles that could guide the design of interactive systems. The authors divide the principles into three main categories, with each category having specific sub-principles that support them. The three main categories, learnability, flexibility and robustness are discussed in the following sub-sections.

4.3.1.1 Learnability

Learnability deals with the ease with which new users can begin effective interaction with an application and achieve maximum performance [Dix et al., 2004]. The specific sub-principles, including their related principles are provided in Table 4.2.

Table 4.2: Principles affecting learnability [Dix et al., 2004]

Principles	**Definition**	**Related principles**
Predictability	Support for the user to determine the effect of future action based on past interaction history.	Operation visibility: refers to how the user is shown the available operations that can be performed next.
Synthesizability	Support for the user to assess the effect of past operations on the current state.	Immediate/eventual honesty: the ability of the interface to provide observable information on changes in the state of the system following user operation.
Familiarity	The extent to which a user's knowledge and experience in other real-world or computer-based domains can be applied when interacting with a new system.	Guessability and affordance: guessability refers to the user's first impression of the system and whether s/he can determine how to initiate any interaction. Affordances are the intrinsic properties of objects that suggest how they can be used.
Generalizability	Support for the user to extend knowledge of specific interaction within and across applications to other similar situations.	
Consistency	Likeness in input-output behaviour arising from similar situations or similar task objectives.	

4.3.1.2 *Flexibility*

Flexibility is the multiplicity of ways in which the end-user and the system exchange information. Dix et al. [2004] identified five specific sub-principles that affect flexibility, together with a number of related principles. These principles are provided in Table 4.3.

Table 4.3: Principles affecting flexibility [Dix et al., 2004]

Principles	Definition	Related principles
Dialogue initiative	Allowing the user freedom from artificial constraints on the input dialogue imposed by the system.	System/user pre-emptiveness: a system pre-emptive system initiates all dialogues and the user simply responds to requests for information. In a user pre-emptive system, the user has the freedom to initiate any action towards the system.
Multi-threading	Ability of the system to support user interaction pertaining to more than one task at a time.	Concurrent vs. interleaved multithreading: concurrent multi-threading allows simultaneous communication of information pertaining to separate tasks Interleaved multi-threading allows temporal overlap between separate tasks but the dialogue is restricted to a single task at any given instant. Multi-modality: allows separate modalities (or channels of communication) to be combined to form a single input or output expression. Multi-modal dialogue may allow concurrent or interleaved modes.
Task migratability	The ability to pass control for the execution of a given task so that it becomes either internalized by the user or the system or shared between them.	
Substitutivity	Allowing equivalent values of input and output to be arbitrarily substituted for each other.	Representation multiplicity: refers to the flexibility for state rendering, for example, using text, sound and graphics. Equal opportunity: blurs the distinction between input and output at the interface.
Customizability	Modifiability of the user interface by the user or the system.	Adaptivity: automatic customization of the interface by the system based on its knowledge of the user. Adaptability: ability of the user to adjust the form of input and output.

4.3.1.3 Robustness

Robustness is the level of support provided to the user in determining successful achievement and assessment of goals [Dix et al., 2004]. The specific sub-principles affecting robustness, including their related principles are provided in Table 4.4.

Table 4.4: Principles affecting robustness [Dix et al., 2004]

Principle	Definition	Related principles
Observability	Ability of the user to evaluate the internal state of the system from its perceivable representation.	Browsability: allows the user to explore the current internal state of the system through the limited view provided at the interface. Static/dynamic defaults: assist the user by passive recall and reducing the number of physical actions necessary to input a value. Static defaults: defined within the system or acquired at initialization. Dynamic defaults: evolve during a session. Reachability: refers to the possibility of navigation through the observable system states. Persistence: the length of time that the effect of a communication lasted and the ability of the user to make use of that effect.
Recoverability	Ability of the user to take corrective action once an error has been recognized.	Forward recovery: involves the acceptance of the current state and negotiation from that state towards the desired state. Backward recovery: attempt to undo the effects of previous interaction to return to a prior state before proceeding. Commensurate effort: stipulates that if it is difficult to undo a given effect on the state, then it should have been difficult to execute in the first place.
Responsiveness	How the user perceives the rate of communication with the system.	Stability: the invariance of the duration for identical or similar computational activity.
Task conformance	The degree to which the system services support all of the tasks the user wishes to perform and in the way that the user understands them.	Task completeness: the level to which the system services can be mapped onto all of the user's tasks. Task adequacy: how the user understands the tasks.

4.3.1.4 Applicability of Dix et al.'s Usability Principles to the Digital Doorway

The three main usability principles of learnability, flexibility and robustness can be applied to the DD.

1. *Learnability:*
   - The DD is based on the assumption of people's ability to independently learn the use of computers; hence the learnability principles of predictability, synthesizability, familiarity, generalizability, and consistency are relevant in this regard.
2. *Flexibility:*
   - With regard to the methods of user input, users should not be forced into a particular method of providing input for the system when other valid methods are possible.
   - Multi-modal presentation of information is also essential in facilitating access to users with varying abilities.
   - Customizability in the form of adaptivity and adaptability are essential to enable both the system and user adjust level of difficulty of game applications based on user performance.
3. *Robustness:*
   - Instantaneous feedback is essential to avoid repeated clicking especially by users who are just learning to use the system. In situations where immediate response is not possible, the user should be informed of progress of the internal processing.
   - Task conformance principle is also relevant as it enables the assessment of the extent to which the system provides appropriate functionalities necessary for the execution of tasks.

Dix et al.'s [2004] principles that are not directly relevant to the DD context are:

- *System versus user pre-emptiveness*: The context of the DD is such that certain user action should be restricted, for instance, changing the system configuration. Hence, the system in this regard mainly initiates all dialogues, so this principle is not included in the list of potential principles for evaluating the DD.
- *Multi-threading*: This principle is not applicable to the DD because the system supports the execution of single task at a time.
- *Task migratability*: This is another principle that is not applicable to the DD context, since the applications that will be evaluated do not require the transfer of control for task execution between the applications and the user.

Although the usability principles by Dix et al. [2004] are comprehensive, they do not provide for the direct accessibility requirements for DD evaluation. The usability principles by Dix et al. [2004] that are relevant to DD evaluation are summarized in Figure 4.3. In Chapter 7, I will integrate these principles with other principles and guidelines to form heuristics for the evaluation of the DD.

1. *Predictability:* Users of interactive systems should be able to determine the effect of future action based on their past interaction history.
2. *Synthesizability:* Users should be able to assess the effect of past operations on the current state of the system.
3. *Familiarity:* Users should be able to apply previous knowledge and experience in other real-world or computer-based domains to a new system.
4. *Generalizability:* Users should be able to extend knowledge of specific interaction within and across applications to other similar situations.
5. *Consistency:* There should be consistency in naming and invocation of similar tasks.
6. *Customizability:* The system or the user should be able to adjust the level of difficulty for educational game applications to suit the user's need.
7. *Observability:* The user should be able to evaluate the internal state of the system based on its perceivable representation.
8. *Multi-modality:* Information should be provided through multiple modes to enable access to people with varying abilities.
9. *Substitutivity:* Whenever appropriate, users should be able to provide input through different input methods.
10. *Responsiveness:* The system's response time should be instantaneous. Whenever this is not possible, the user should be informed of the execution progress.
11. *Task conformance:* The system should provide the functionalities that enable users to execute the tasks required to achieve their goals.

Figure 4.3: Dix et al.'s [2004] principles relevant to Digital Doorway context

4.3.2 Gelderblom's Guidelines for the Design of Children's Technology

Gelderblom [2008] developed an extensive set of guidelines for the design of technology for children aged five to eight, based on: the study of psychological development of children and the way it impact on their use of technology; and, existing guidelines and principles for the design of children's application.

The 350 guidelines that emerged from the study were grouped into six categories and 26 sub-categories. The six main categories are:

1. Guidelines to ensure the developmental appropriateness of applications.
2. Guidelines aimed at the development of specific skills in children, for example, reading and writing skills.
3. Guidelines on the design of built-in support.
4. Guidelines aimed at encouraging collaborative use of technology.

5. Guidelines on addressing user diversity.
6. Guidelines on the use and design of interactive environments and devices.

While acknowledging that these guidelines were formulated with younger children in mind, some of them can, nevertheless, be applied to all age groups and users with disabilities. Section 4.3.2.1 describes Gelderblom's [2008] guidelines that could be used to derive heuristics for evaluating the DD.

4.3.2.1 Applicability of Gelderblom's Guidelines for the Development of Children's Technology to the Digital Doorway

In deciding whether or not to include specific guidelines from the six categories, their level of generality was used as the criterion for inclusion. For example, the guideline "Reduce the cognitive load required for interaction so that there are sufficient cognitive resources for learning" (based on theories on the development of working memory and the development of computer literacy skills by children) was found to be general enough to be applicable to users of all age group. Hence this guideline was included in the list of potential guidelines that could be used to formulate heuristics for evaluating the DD. As another example, the guideline "Design technologies to reflect a child's context. Applications should ideally, come in multiple languages, reflecting gender equity and avoiding racial discrimination" was found to be a useful guideline to assess the extent to which the applications evaluated reflects the diversity of the different communities where DDs are deployed. Gelderblom's [2008] guidelines which are relevant to my study are summarized in Figure 4.4.

1. Reduce the cognitive load required for interaction so that there are sufficient cognitive resources for learning.
2. Users should not be forced to remember instructions or previous choices. All objects, actions, and options should be clearly visible.
3. Children should be able to apply their real-world or other computer-based knowledge when interacting with a new system.
4. Successful interpretation of an icon depends on its caption. Icons and interface elements should be given meaningful and intuitive captioning so that users can successfully interpret their functionalities.
5. There should be a clear mapping between interface elements and their effect on the system.
6. Children should be able to determine the effect of future action based on past interaction history.
7. Provide adequate feedback in the form of audio, tactile, verbal or visual information about user action and their effect.
8. Provide immediate feedback to blind users, for example, through subtle sound feedback.
9. Appropriate hints should be provided to enable children correct cognitive mistakes.
10. Children should not be forced to remember audio instructions.

11. Keep complexity levels low for beginners, but provide a high enough ceiling to allow users of different levels to benefit.
12. Unintended user errors should be prevented through the use of constraints at strategic points.
13. The system's response time should be instantaneous. Whenever this is not possible, the user should be informed of the execution progress to prevent repeated clicking or hitting of keys.
14. Interactive systems should reflect a child's context and provide multiple languages from which children can select. The system should reflect gender equity and avoid racial discrimination.

Figure 4.4: Gelderblom's [2008]guidelines applicable to Digital Doorway context

4.3.3 Nielsen's Heuristics

In 1990, Nielsen and Molich proposed nine heuristics for evaluating the usability of user interfaces. The heuristics were revised in 1994 by Nielsen based on analysis of a database of 249 usability problems. The revised version of Nielsen's heuristics are reproduced below [Nielsen, 1994b]:

- *Visibility of system status:* The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
- *Match between system and the real world:* The system should speak the user's language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.
- *User control and freedom:* Users often choose system functions by mistake and will need a clearly marked 'emergency exit' to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.
- *Consistency and standards:* Users should not have to wonder whether different words, situations or actions mean the same thing. Follow platform conventions.
- *Error prevention:* Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.
- *Recognition rather than recall:* Minimize the user's memory load by making objects, actions and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.
- *Flexibility and efficiency of use:* Accelerators, unseen by the novice user, may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.
- *Aesthetic and minimalist design:* Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

- *Help users recognize, diagnose and recover from errors:* Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.
- *Help and documentation:* Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

4.3.3.1 Applicability of Nielsen's Heuristics to the Digital Doorway

Nielsen's [1994b] heuristics overlap with the usability principles by Dix et al. [2004], discussed in section 4.3.1. The self-teaching nature of the DD means that users should be able to use the system without referring to documentation. However, help in a contextualized form should be available so that users can access them when required.

All Nielsen's [1994b] heuristics, except one, were found to be applicable to the DD context. The heuristic which is not directly applicable to the evaluation of the DD:

- *Flexibility and efficiency of use:* As stated in section 2.4.1, the DD keyboard does not provide function and special keys, such as, the *ctrl* and *Alt* keys, which are required for initiating shortcut commands.

Figure 4.5 provides a summary of Nielsen's [1994b] heuristics relevant to the DD environment.

1. *Visibility of system status:* Appropriate feedback should be provided to the user within reasonable time.
2. *Match between system and the real world:* Language use should be those that are familiar to the user. Technical terms must be avoided.
3. *User control and freedom:* Users should be able to reverse unintended actions.
4. *Consistency and standards:* There should be consistency in the naming and invocation of similar tasks. Industry and platform conventions should be followed.
5. *Error prevention:* Applications should be designed such that errors are prevented from occurring.
6. *Recognition rather than recall:* Short-term memory load should be minimized. All objects, actions, and options should be visible. The user should not be forced to recall information from one screen to another.
7. *Aesthetic and minimalist design:* The user interface should not be cluttered with irrelevant information, control buttons and icons.
8. *Help users recognize, diagnose and recover from errors:* Error messages should be expressed using simple terms. The location of problems should be made obvious and ways of correcting them provided.
9. *Help and documentation:* Help documents should be easy to access and focused on the user's task.

Figure 4.5: Nielsen's heuristics relevant to Digital Doorway context

4.3.4 Usability and User Experience Goals of Preece, Rogers and Sharp

Before beginning the design of an interactive product, designers should consider the primary objective of the product they wish to design. An application aimed at supporting rapid completion of tasks, for instance, should be highly efficient. According to Preece et al. [2007], classifying these objectives into usability and user experience goals can help designers to identify design objectives. Usability goals are concerned with meeting specific usability quality criteria while user experience goals have to do with the way users feel about the product. Although it is not possible to design *a* user experience, designers should strive to design *for* a user experience [Preece et al., 2007].

4.3.4.1 Usability Goals

Preece et al. [2007] identified the following six usability goals:

1. *Effectiveness* a general goal referring to how good a product is at doing what it is supposed to do, that is, how well a user can achieve his/her goal using the product.
2. *Efficiency* is the way a product supports users in carrying out their tasks, that is, how fast users can complete a given task using the product.
3. *Safety* deals with protecting users from dangerous conditions and undesirable situations, that is, the built-in mechanisms that are provided to reduce inadvertent errors by users.
4. *Utility* is the extent to which the product provides the right kind of functionality so that users can do what they need or want to do.
5. *Learnability* is concerned with the ease of learning to use the system.
6. *Memorability* is the ease with which users can remember the use of the system, after a period of not using it.

4.3.4.2 User Experience Goals

User experience refers to how users feel about a product. While usability goals are used to objectively assess a product in its own right, user experience goals are more subjective in that they are based on the feelings of individual users. In the past, HCI practitioners have only focused on usability goals, such as, effectiveness and learnability. In resent times, there is increase awareness about features that evoke user experience.

User experience can be positive or negative. For example, products may be designed to be enjoyable, satisfying, pleasurable, motivating, challenging, aesthetically pleasing, boring, or annoying. Usability and user experience goals may conflict with one another; the primary purpose of the application should guide designers in making appropriate trade-offs [Preece et al., 2007].

4.3.4.3 Applicability of Preece et al.'s Usability and User Experience Goals to the Digital Doorway

The six usability goals identified by Preece et al. [2007] are similar in their meaning to the usability principles by Dix et al. [2004] and are all relevant to the evaluation of the DD. Although the DD is not a transaction processing system, where task completion time is an important measure of usability, application efficiency in the form of quick response time is important. A slow response often result in repeated clicks, especially from novice users. This may interfere with the execution of applications. User experience goals such as motivation and challenge are relevant as they are essential elements of effective educational games.

4.3.5 Norman’s Principles of Design

Norman [2001] identified certain basic principles that guide the design of products, including interactive systems. These principles enable users to determine what should be done to complete a task based on the characteristics of the object. The five design principles identified by Norman [2001] are:

1. *Visibility:* Visible controls and functions enable users to determine what needs to be done next. Adding more functionality to an application sometimes result in designers ‘hiding’ the controls for the functions to avoid cluttering the interface.
2. *Mapping:* This is the relationship between controls and their results in the world. Exploiting the physical characteristics of objects or cultural standards can facilitate learning how to use them. For instance, the use of *down* and *up* arrow keys on the keyboard to move the cursor down and up in a text editor program respectively capitalizes on the mapping of these activities to the keys.
3. *Feedback:* The provision of information to the user regarding what action has been taken and the effect of that action. Feedback is an essential principle of design as it enables users to evaluate whether their goals have been met and to determine further action. Lack of feedback could result in a user invoking the same command repeatedly, thinking that earlier commands were ineffective.
4. *Constraints:* These are the mechanisms that restrict allowable actions at specific time. One example of constraint in an interactive computer system involves ‘greying-out’ certain disallowed options.
5. *Affordances:* These are attributes of an object which enable people to determine how to use it. Taking advantage of the affordances of an object provides users with clues as to how to use and operate it.

4.3.5.1 Applicability of Norman’s Design Principles to the Digital Doorway

Four of the five principles by Norman [2001] are directly applicable to DD context. The principle of visibility and feedback were discussed earlier as being relevant to the evaluation of the DD. Exploiting natural mappings and building in constraints at strategic places are essential principles that could enhance the usability of the DD.

The principle of affordance has been excluded because it was difficult to formulate a heuristic that could be easily used to assess the conformance of interface elements to the principle.

Norman’s [2001] design principles that can guide the development of evaluation heuristics for the DD are summarized in Figure 4.6.

1. *Visibility:* All systems' controls and functionalities should be visible so users can determine the next required actions.
2. *Mapping:* The mappings between controls and their effect should be intuitive and easily understood.
3. *Feedback:* Adequate information should be provided to the user regarding what action has been taken and the effect of that action.
4. *Constraints:* Unintended user errors should be prevented through the use of constraints at strategic points.

Figure 4.6: Norman's design principles relevant to Digital Doorway context

4.3.6 Shneiderman's Golden Rules of Interface Design

Shneiderman [1998] provides eight design principles called 'Golden Rules' for the design of interactive systems. These golden rules can be used at the design stage and as a checklist to evaluate an interactive system. The eight golden rules by Shneiderman's [1998] are reproduced below:

1. *Strive for consistency:* Consistency is one of the most important design rules and the most frequently violated. There should be consistency in the type of action sequences required for similar situations. There should be consistency in interface layout and terminology use within and across applications. Whenever there is a valid reason to deviate from this rule, then this should be comprehensible to users.
2. *Enable frequent users to use shortcuts:* Users who have become proficient in the use of the system following repeated use should be able to bypass dialogues that are not pertinent to the completion of the task at hand.
3. *Offer informative feedback:* The system should provide appropriate feedback for every user action. Feedback should be comprehensive, taking into account users' age and experience. Response time should be commensurate to the magnitude of the action.
4. *Design dialogues to yield closure:* Appropriate feedback should be provided so that users know when they have completed a task.
5. *Offer error prevention and simple error handling:* Systems should be designed such that it becomes difficult for users to make serious mistakes. When users do make mistakes, they should be provided with clear and understandable instructions to enable them to recover.
6. *Permit easy reversal of actions:* It should be possible for users to undo a previous action. This helps to relieve anxiety and encourage system exploration since users know that they reverse an unintended consequence.
7. *Support internal locus of control:* Systems should be designed to allow the user to initiate actions as much as possible instead of him/her having to respond to the system's actions all the time.

8. *Reduce short-term memory load:* Interface elements and displays should be kept simple since there is limit to the amount of information that can be held in working memory at any time. Multiple page displays should be consolidated and adequate time allowed for learning codes, mnemonics and action sequences.

4.3.6.1 Applicability of Shneiderman's Golden Rules to the Digital Doorway

Most of Shneiderman's [1998] golden rules are similar in their meanings and interpretations to the design and usability principles discussed thus far. The principles of consistency, feedback, error prevention, and support for object recognition (as opposed to recall) are all useful in the formulation of evaluation heuristics for the DD. Golden rule four "Design dialogues to yield closure" has been interpreted as being similar in its meaning to the provision of informative feedback in rule three.

The following golden rules by Shneiderman [1998] are excluded from the list of guidelines that could be used to formulate heuristics for evaluating the DD:

- *Enable frequent users to use shortcuts*: As stated in section 2.4.1, the DD keyboard does not provide the special keys necessary for initiating commands through shortcuts; golden rule two is therefore excluded.
- *Support internal locus of control*: The amount of actions that can be initiated by users of the DD is quite limited (see section 4.3.1.4), hence golden rule seven has been excluded.

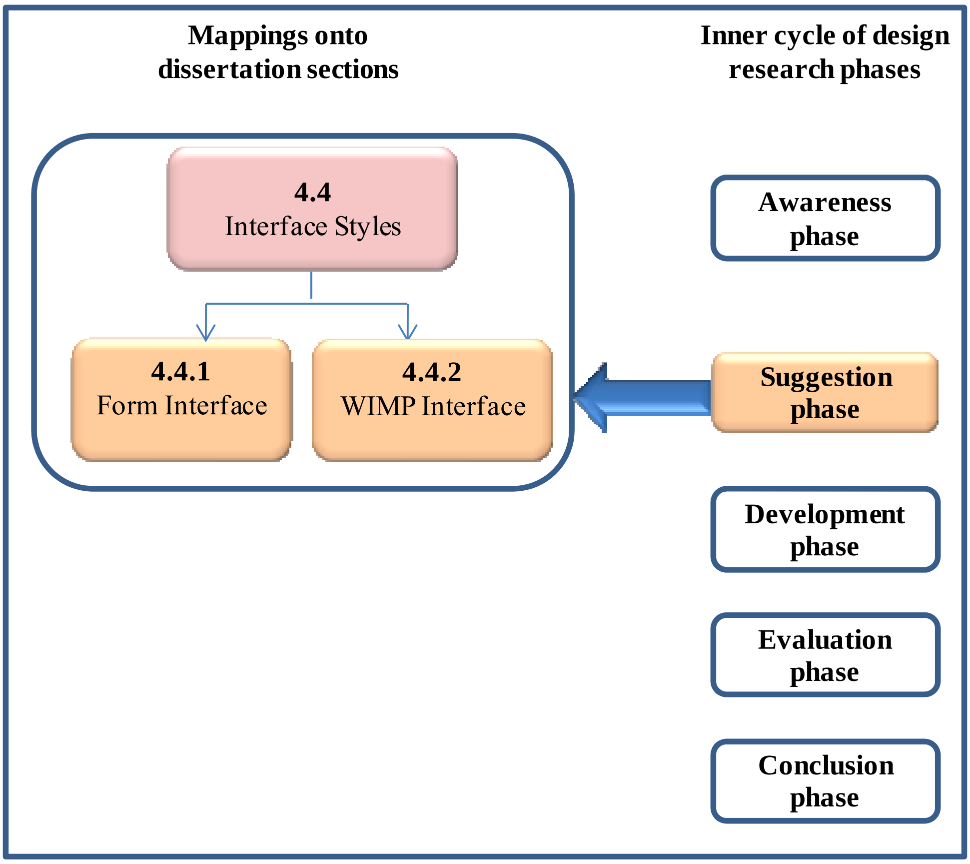
Shneiderman's [1998] golden rules that are relevant to the DD context are summarized in Figure 4.7. In Chapter 7, these golden rules and the other principles discussed in section 4.3 will be integrated to develop the heuristics for usability evaluation of the DD.

1. *Strive for consistency:* There should be consistency in the naming and invocation of similar tasks
2. *Offer informative feedback:* Appropriate information should be provided to the user regarding what action has been taken and the effect of that action taking into consideration users' age and experience. Such feedback should be provided within reasonable time.
3. *Offer error prevention and simple error handling:* Unintended user errors should be prevented through the use of constraints at strategic points. When errors do occur, users should be provided with clear and understandable instructions on how to recover the error.
4. *Permit easy reversal of actions:* Users should be able to retract previous and unintended actions to facilitate system exploration.
5. *Reduce short-term memory load:* Short-term memory load should be minimized. All objects, actions, and options should be visible. The user should not be forced to recall information from one screen to another.

Figure 4.7: Shneiderman's golden rules applicable to Digital Doorway context

4.4 INTERFACE STYLES

Mapping of section 4.4 to the design research phases



Interaction is the communication that occurs between the user and the computer (see section 2.2). The choice of interface used in a system can impact on the effectiveness of the interaction. The two main interface styles in the DD are: the electronic form; and, the windows, icons, menus and pointers (WIMP) interface. This section briefly describes these two interface styles, and the principles guiding the design of a form interface.

4.4.1 Form Interface

Form interfaces are typically based on a paper form. Forms include a number of fields where the user is expected to type the required information and each field has a label or caption specifying the type of data required. Well-designed form interfaces are easy to learn and use. They are also flexible, allowing easy movement around the form [Dix et al., 2004].

4.4.1.1 Design Principles and Guidelines for Form Interfaces

The DD utilizes a simple form interface, described in section 8.2.2, to collect user information from those who wish to use the system as registered users. It is therefore essential to examine the principles that guide the design of form interfaces. Mayhew [1992] provides six categories of design principles that should be considered when designing a form interface. The principles are [Mayhew, 1992]:

1. *Form organization and layout:* The guidelines within this category are concerned with the organization of screen elements and the layout of form fields.

    - *Design and organize form to support tasks:* The organization and layout of form fields should adequately support the user's tasks. When data is captured from a paper-based form the layout, organization and fields of the electronic form should be similar to the paper form. When data can be input in any order, cursor movements should be flexible to support this form of data entry.
    - *Organize groups of items related semantically:* Form elements should be grouped together based on their semantic relationship. Elements can be grouped according to the sequence of use, especially when users will be working from a paper-based form. An alternative grouping approach is the frequency of use of field elements.
    - *Separate logical groups using visual cues:* Logical groupings should be separated using spaces, lines, colours or other visual cues. The use of colour to separate logical grouping has advantage as it does not take up additional screen space.
    - *Keep related and interdependent items on the same screen*: Fields that provide context to other fields should be located on the same screen. Users should not be required to recall information from one screen to another.

2. *Form caption and field design:* These guidelines are related to the design of form fields and their captions.

    - *Captions and fields should be justified based on application users, their tasks and data type:* Captions should be left justified if they are of similar lengths and users will not input data from a source document but if caption lengths vary significantly, then the data fields should be left justified.
    - *Provide distinctive field group and section headings in complex forms:* When a form has many data fields, the fields should be grouped together based on their semantic relationship using spaces and borders. In addition, descriptive names should be given to each group to facilitate ease of searching specific fields.
    - *Distinguish captions from fields:* Form captions should be distinguished from data fields through visual clues, such as, bold face and colour coding. Data fields should be made more prominent than captions since fields are the focus of attention.
    - *Captions should be brief, familiar and descriptive:* The use of wordy captions should be avoided as they occupy more screen space. Avoid the use of technical terms that are not comprehensible to users. A caption should be descriptive of the field it represents. Abbreviations should only be used if they are well known and standard in the user domain.
    - *Indicate the number of character spaces available in a field:* Providing users with an indication of maximum allowable characters before they begin filling in a field guides the decision on a more concise and less ambiguous input.

- *Indicate when fields are optional:* Forms typically include fields that must be filled and those that may be left unfilled. Clearly marking optional fields that may be left blank will reduce unnecessary keystrokes.

3. *Form input format:* The guidelines within this category are aimed at speeding up user input and reducing input errors.

- *Consider providing system completion of unambiguous partial input:* Automatic completion of certain predetermined, unambiguous entries will facilitate ease of entry and reading of form fields. For example, when a user enters 'Ja' or '1' for 'January' and 'Jun' or '6' for 'June' and the cursor is placed on the next field, the system should complete the entry.
- *Consider providing pop-up or pull-down menus for form fields with many but well-defined entry options:* When the application use is infrequent, when users are inexperienced, when the number of valid inputs is large, and/or when inputs are difficult to spell or remember, it may be beneficial to provide pop-up or pull-down menus with fields to present entry options. Fields with entry menus should be distinguished from those without through visual clues.
- *Provide defaults whenever possible. Allow simple acceptance of defaults:* Where there is a most likely or a most commonly entered value, the field should be pre-filled with the value to reduce keystrokes and typing errors. The application should allow the acceptance of the default value using a simple key like the 'Tab' key.
- *Make high-frequency inputs easy to express:* The number of keystrokes for high-frequency users and/or high-frequency inputs should be minimized while at the same time maximizing the ease of learning and memorability for occasional users and/or low frequency inputs. For example, if the form contains many yes/no questions, a lowercase 'y' or 'n', which are faster to enter, should be acceptable in addition to uppercase letters or the full words.
- *Data input should be meaningful to domain users:* Data or codes should be based on common English usage as these will be easier to learn and remember than arbitrary codes. For example, the use of 'M' and 'S' for 'Married' and 'Single' respectively.
- *The application should be 'case blind' when it does not really matter:* In situations where the meaning of data is not affected by its case, for example 'YES' or 'Yes' or 'yes', the system should recognize as correct any of the versions and not force users to enter any one.

4. *Form prompts and instructions:* The guidelines in this category are aimed at improving the design of prompts and instructions for completing the form.

- *Prompts should be provided to guide users:* When forms are completed by casual users, when user input must be in specific formats or when users are not entering data from a source document, prompts should be provided to guide and remind users of the type of information required.

- *Prompts should be brief and unambiguous:* Prompts should be provided in simple, concise and clear terms without sacrificing important information, for example, the range of allowable values.
- *Prompts should be placed to the right of data fields or in a MicroHelp line at the bottom of the screen:* Provide prompts through MicroHelp (a contextualized help tips associated with the currently active field) at the bottom of the screen to save screen space. Placing prompts below data fields could confuse users as to which field the prompt applies to when the form is crowded. Prompts that are in-between captions and data fields force frequent users who no longer require the prompt to read them.
- *Instructions should be placed at consistent locations across screens:* When the form is distributed across several screens, instructions should be at the same regions of all screens. The instructions should be easily distinguishable from other form elements.

5. *Form Navigation:* The guidelines within this category are aimed at ensuring easy movement within and across form elements.
    - *When the form is first entered, place the cursor in the most likely default field:* In forms that require most or all the fields to the completed, the cursor should be placed in the first data field. When only a few fields of the form are to be completed the cursor should be positioned in the most likely default place.
    - *Allow backward and forward movement by field and within fields:* Users should be able to move backward and forward within a field using the keyboard to enable input edition rather than having to retype the whole field.
    - *Provide screen titles and page numbers on screens in multi-screen forms:* Including screen title and page numbers on the screens of multiple screen forms facilitates user orientation and enables them to determine the number of screens that have been completed and those that still have to be filled.
6. *Form error handling:* These guidelines are concerned with the ways to deal with unavoidable user errors when completing forms.
    - *Allow character edits in fields:* The form should allow users to correct typing errors character by character and should not force them to erase and retype the entire field.
    - *Place the cursor in the field where the error occurred:* After detecting user input error, the cursor should be positioned in the field where the error has occurred, with the error field highlighted. This will allow the user to easily detect and correct the error.
    - *Provide error messages that are meaningful to users based on their knowledge:* Error messages should be given in simple and comprehensible terms taking cognisance of the application domain of use and users' knowledge.

4.4.1.2 *Application of Mayhew's guidelines to the Digital Doorway*

All the six guideline categories, proposed by Mayhew [1992] can be applied to the DD. However, not all the sub-categories are relevant to the DD context. Examples of inapplicable guidelines are those relating to the provision of default values for data fields, auto-completion of user entry, and giving prompts through MicroHelp. These guidelines are aimed at the design of more complex electronic forms. Only nine out of the potential 26 guidelines are applicable to the DD context. The guidelines that are relevant to DD context are summarized in Figure 4.8.

1. *Organize groups of items related semantically:* Form elements should be grouped together based on their semantic relationship.
2. *Separate logical groups using visual cues:* Provide visual reinforcement for groups of elements through efficient use of white spaces and borders.
3. *Captions should be brief, familiar and descriptive:* Give meaningful names to field captions.
4. *Indicate the number of character spaces available in a field:* Clearly specify the limit for data having minimum or maximum allowable length.
5. *Indicate when fields are optional:* Designate required fields in standard and consistent ways.
6. *When the form is first entered, place the cursor in the most likely default field:* Provide visible cue by positioning the cursor in the first field at the start of the form.
7. *Allow backward and forward movement by field and within fields:* Users should be able to edit data fields by moving the cursor backward and forward rather than having to retype the entire field.
8. *Place the cursor in the field where the error occurred:* When input errors are detected, the cursor should be positioned in the error field, highlighted.
9. *Provide error messages that are meaningful to users based on their knowledge:* Give feedback for missing data using clear and unambiguous terms, taking into account the user's age and experience.

Figure 4.8: Mayhew's guidelines for form interface applicable to Digital Doorway context

4.4.2 WIMP Interface

A *WIMP* interface consists of four key elements: windows; icons; menus; and pointers. Many of the contemporary interactive systems utilize *WIMP* interfaces. Apart from the four key elements, other components of a *WIMP* interface are buttons, toolbars, palettes and dialogue boxes.

A window is a screen region consisting of text and graphics, which can be moved or resized. Other elements associated with windows are scrollbars, toolbars and menu bars that enhance their utility. For example, scrollbars enable the user to move a window's content upward, downward and sideways. Typically, more than one window can be on the screen at the same time, thus enabling the visibility of multiple tasks.

An icon is a miniature image representation of system elements, such as, a window, a file or a printer. Icons enable quick and easy access to system resources by clicking on them. An icon can be a realistic representation of an object or an arbitrary symbol. Users may find it difficult to recognize arbitrary symbols [Dix et al., 2004; Mayhew, 1992].

A menu provides a list of available options and its structure of can be hierarchical, linear or networked. Linear menu interfaces support learnability through the visibility of available options because users do not have to recall how to use the system but hierarchical and networked menus can be problematic because the required option may not be at the top layer of the hierarchy, thus requiring the user to wade through a number of levels [Dix et al., 2004; Mayhew, 1992].

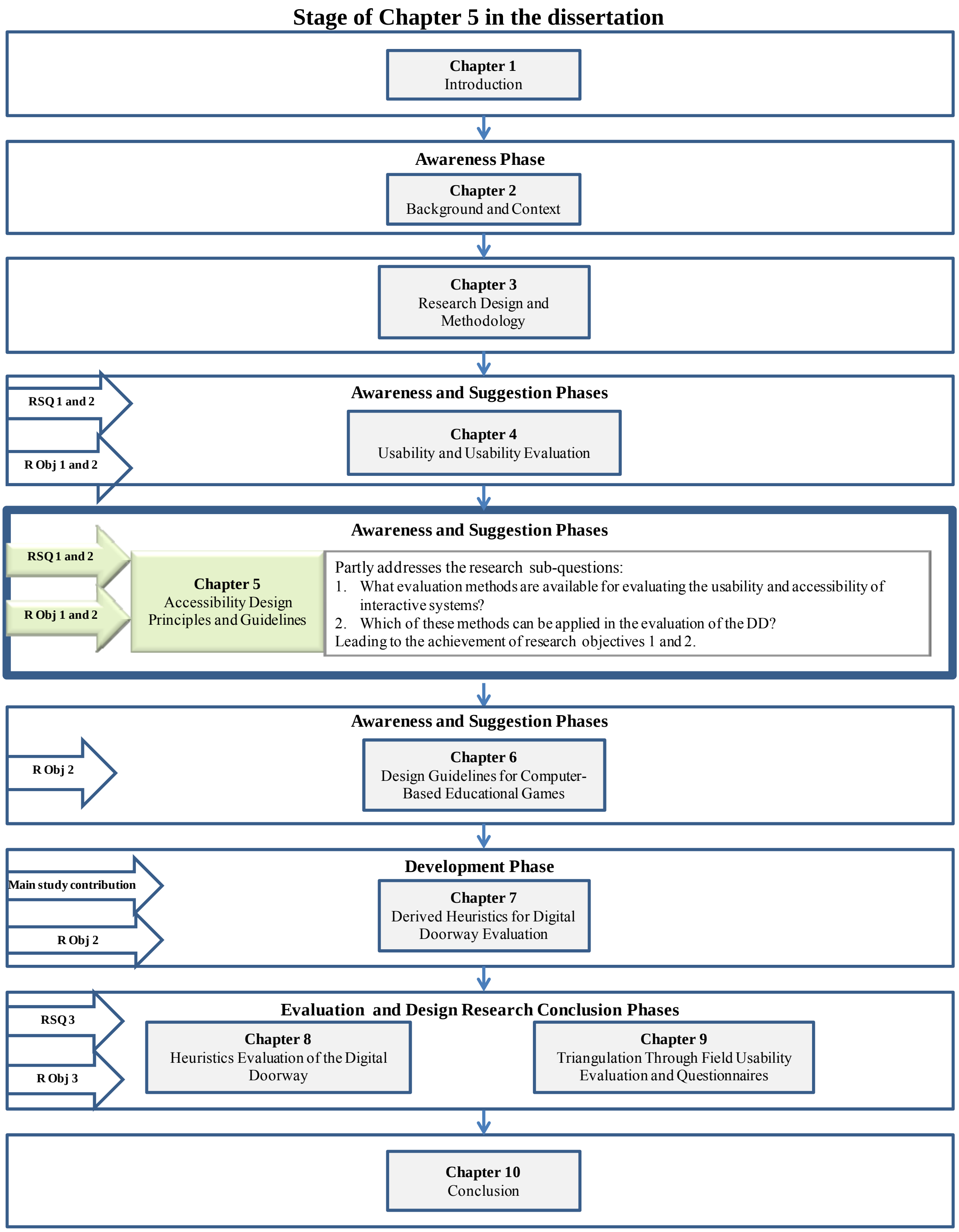
The fourth key element in a *WIMP* interface is the pointer. Using a pointing device such as a mouse or trackball, the user can select system's elements, such as, an icon. A pointer can have different cursor shapes to distinguish system modes. Pointers also have hot-spots that indicate where they point at a particular time [Dix et al., 2004].

The principles and guidelines for the design of *WIMP* interfaces are not examined in this dissertation, sine none of the authors of the design principles, guidelines and heuristics reviewed in section 4.3 explicitly provide guidelines for the design of *WIMP* interfaces, neither has Mayhew [1992]. However, the usability of *WIMP* interface can be adequately covered by the principles, guidelines and heuristics already discussed in section 4.3.

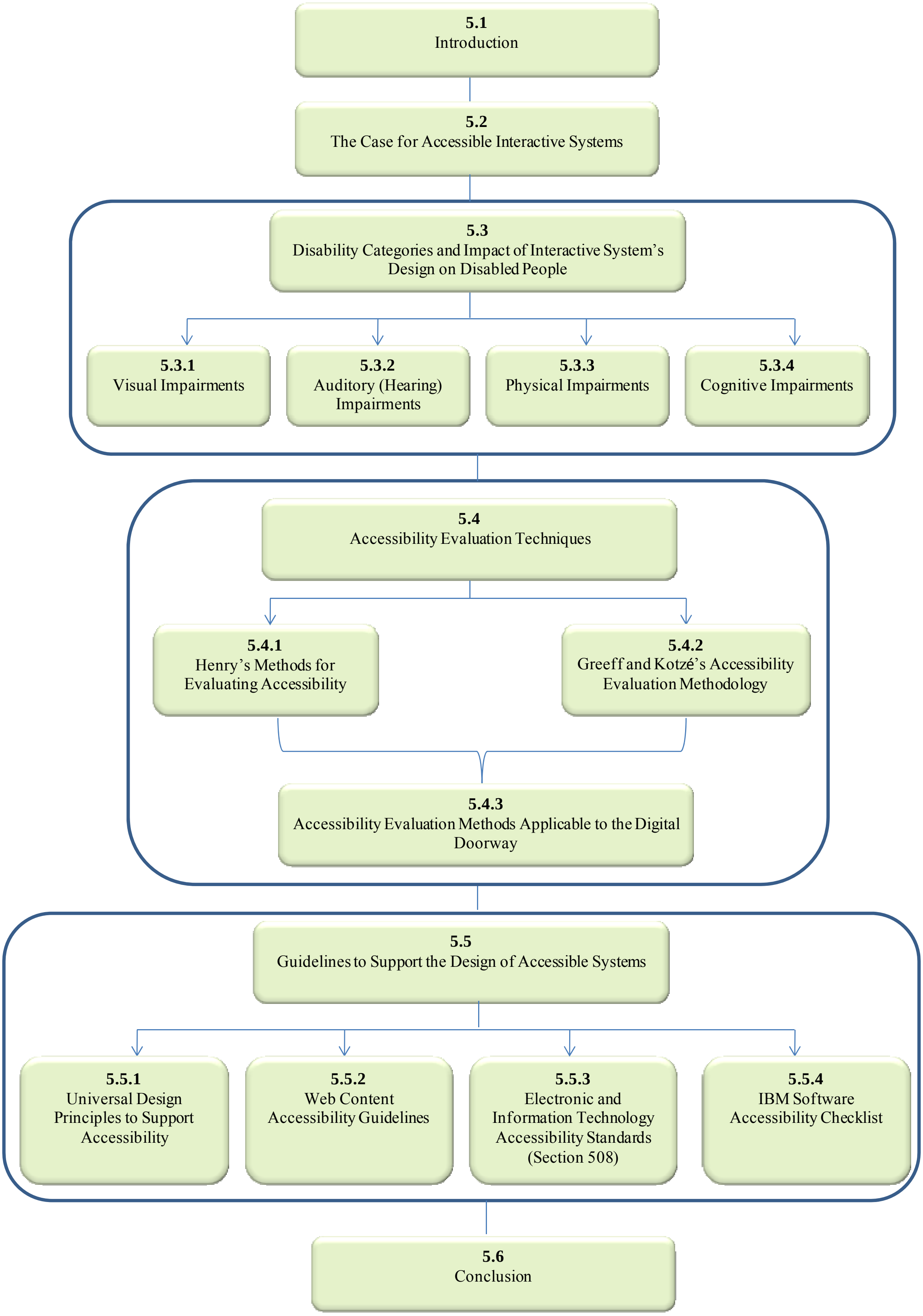
4.5 CONCLUSION

In this chapter, I have provided detailed discussion of different usability evaluation methods, their benefits and limitations and the methods that are appropriate for evaluating the usability of a selection of interfaces and applications installed on the DD were identified. I also looked at the design principles and guidelines proposed by various authors to support the usability of interactive systems with the aim of determining their applicability to the DD context. Looking ahead, in Chapter 5, I will discuss accessibility and the principles guiding the design of accessible systems.

CHAPTER 5: ACCESSIBILITY DESIGN PRINCIPLES AND GUIDELINES



Map of Chapter 5



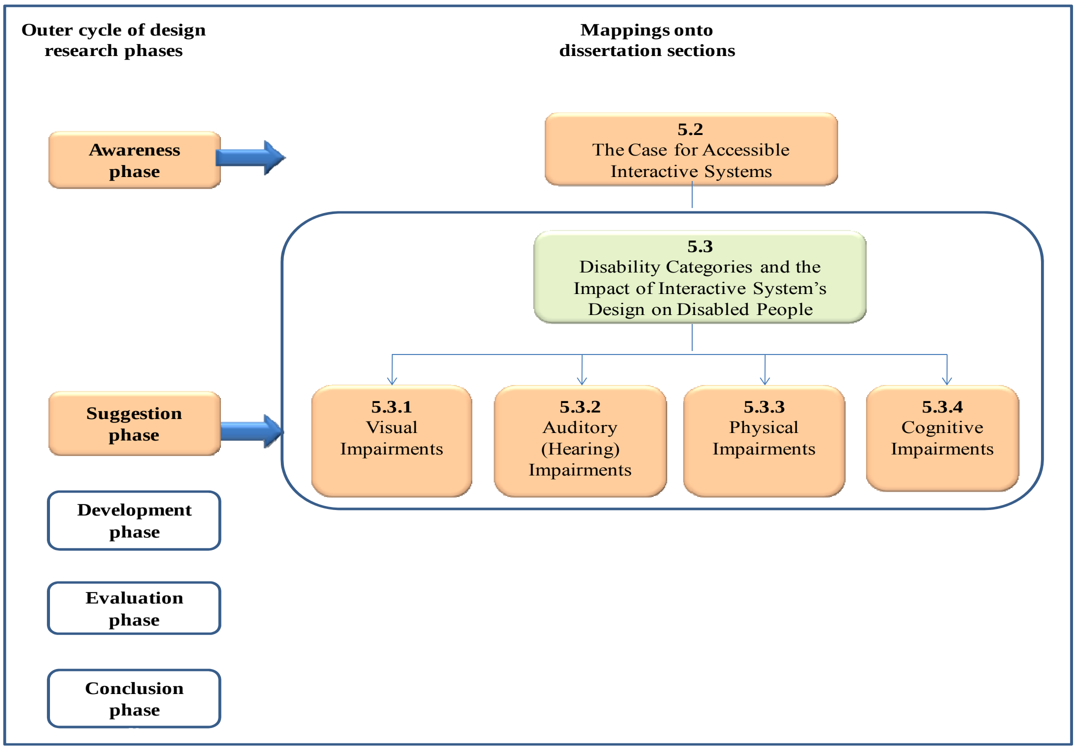
5.1 INTRODUCTION

This chapter also provides part of the answers to my research sub-question 1: What evaluation methods are available for evaluating the usability and accessibility of interactive systems? and sub-question 2: Which of these methods can be applied in the evaluation of the DD? It is the second of three literature investigation chapters representing the suggestion phase for the outer design research cycle, and the awareness and suggestion phases of the inner cycle of my design research.

I start by providing reasons for providing accessibility support in interactive systems (section 5.2). In section 5.3, I examine the four disability categories and their implications for the design of interactive systems. To guide the decision on the appropriate methods for evaluating direct accessibility support in the DD, I look at the methods for evaluating the accessibility of interactive systems in section 5.4. Specifically, I review the accessibility evaluation methods of Henry [2007], and the lightweight accessibility evaluation methodology of Greeff and Kotzé [2009]. In section 5.5, I discuss guidelines for the design of accessible system, starting with the universal design principles. This is followed by the discussion of three sets of accessibility guidelines: Web Content Accessibility Guidelines [1999]; the United States' electronic and information technology accessibility standards [2000]; and, IBM software accessibility checklists [2009]. The guidelines are examined for their appropriateness in the development of heuristics to evaluate direct accessibility support in the DD. I conclude the chapter in section 5.6.

5.2 THE CASE FOR ACCESSIBLE INTERACTIVE SYSTEMS

Mapping of sections 5.2 and 5.3 to the design research phases



This section maps onto the awareness phase of the outer cycle of the design research by providing some of the reasons why designers should be concerned with the accessibility of the systems and applications they develop. Although the primary goal of accessibility is to make interactive systems usable for people with disability, an accessible system is also beneficial to people without disabilities, where the environment of usage imposes limitations in the use of an application. For example, in a very quiet room where noise is prohibited, text-based feedback will be more appropriate than audio feedback. This kind of constraint is referred to as *situational limitations* [Henry, 2002].

Many countries of the world have a considerable number of people with various forms of ability limitations, often referred to as disabilities. In the United States, a 2008 community survey revealed that more than 36 million people were disabled at the time [Erikson, Lee and von Schrader, 2010]. A 2007 community survey conducted by Statistics South Africa showed that nearly two million people, accounting for four percent of the total population had some form of disabilities [Statistics South Africa, 2007]. Furthermore, the United Nations [2006] estimated that more than 500 million people around the world are disabled. These statistics show that there are large number of people with disabilities around the world and provide a compelling reason to address accessibility concerns.

Henry [2002] provides a number of motivations for incorporating accessibility into the design of Web sites. These motivators are also applicable to non Web-based interactive systems:

- *Compliance with regulatory and legal requirements:* The primary reason why organizations are addressing accessibility issues is to ensure compliance with legal and government regulations. In 1999, an Australian blind user successfully sued the Sydney organizing committee for the Olympic games under the Australian disability discrimination Act (DDA) because he was unable to order game tickets using Braille technology [Waddell, 2002]. The United State Government's Section 508 of the Rehabilitation Act stipulates that all federal electronic information should be accessible to people with auditory, visual and mobility impairments.
- *Exposure to more people:* Advances in health care technologies and living standards, especially in developed countries, have resulted in growth of the elderly population. This means that many people are now living longer. A substantial number of this population group will develop some form of degenerative disability due to advanced age [Darzentas and Miesenberger, 2005]. The stereotype that senior citizens are averse to the use of new technologies is not necessarily true [Dix et al., 2004]. Senior citizens and disabled people have more compelling reasons to use new technologies, provided special concerns that impact on successful access are addressed. For example, people who are unable to drive, those with mobility difficulty, or difficulty carrying packages, are more likely to find accessible online shopping as an attractive alternative to the traditional 'bricks and mortar' shops. Furthermore, communication technologies like e-mail and instant messaging can provide social interaction to those restricted by mobility and speech impairments.

In addition, the disabled and the elderly constitute a viable target market for organizations due to their powerful impact on the overall economy. It is estimated that disabled people in the United States have a combined annual income of over $175 billion [Burks and Waddell, 2001].

- *Exposure to more situations:* Situational limitations involve circumstances affecting a particular application or the prevailing environment of use. Situational limitations that affect both disabled and able persons can influence the design of interactive systems. Such limitations might be inherent in the application itself, for example, personal digital assistants (PDAs) and mobile phones, frequently used to access the Internet, have limited input/output capabilities. Compliance with accessibility guidelines can optimize interaction with these devices. An example of situational limitation involving the environment of usage relates to the use of an application in a noisy environment. The provision of information through sound alone in this environment will not be appropriate.
- *Better design and implementation:* Incorporating accessibility into design results in better design and implementation. In addition to enabling access to the disabled, accessibility improvement also enhance usability for users without disability.
- *Cost savings:* Although the initial cost of incorporating accessibility features into a design might be high, the overall development and maintenance costs are reduced. An accessible e-commerce site will result in more sales because more people will be able to access the site. Addressing accessibility issues will also reduce legal expenses that could result from lawsuits by users who might want to enforce their right to equal treatment.
- *Enhancement of corporate reputation:* An organization that is committed to accessibility will be viewed in a positive light by the community in which it operates. An organization that is compliant with accessibility standards and guidelines can also include this information in marketing materials so that potential users can be aware of this. For example, a Web site that conforms to the WCAG 1.0 [1999] can include a logo indicating such compliance on its Web site.
- *Enlightened self-interest:* When organizations are aware of the benefits of accessibility, they can proactively do something to address accessibility issues so that they can reap the benefits. This is known as 'enlightened self-interest'.

Although it has been said that legal concerns is the primary reason why organizations address accessibility concerns, ethical consideration is also an important factor. Every citizen should have equal opportunities, particularly with regard to the right to education and employment opportunities. Without proper education, the probability of being gainfully employed is reduced. Increasingly, the Web is becoming an essential portal to access and share information. The exclusion of certain people from this platform because of their ability is morally and ethically incorrect [Darzentas and Miesenberger, 2005].

The choices that are made during the design of interactive systems could impact on the ability of people with disabilities to effectively access the system. Section 5.3 will therefore examine the different types of disabilities, the types of limitations associated with each disability, and how these users are affected when using interactive systems.

5.3 DISABILITY CATEGORIES AND IMPACT OF INTERACTIVE SYSTEMS' DESIGN ON DISABLED PEOPLE

In general, disabilities are classified as visual, auditory (hearing), physical or cognitive. This section examines these disability categories and the effect of interactive systems' design on users with disabilities.

5.3.1 Visual Impairments

5.3.1.1 Colour Blindness

Colour is composed of three components: hue; saturation; and, lightness.

1. *Colour hue* is a visual sensation that varies according to light wavelengths. Hue enables the identification of basic colours [Mayhew, 1992]. The colour hue circle with eight different colours is depicted in Figure 5.1.
2. *Saturation* is a sensation that is related to the number of different wavelengths, and contributes to the sensation of colour. Typically, the human eyes do not perceive a pure set of single wavelength corresponding to a pure colour but several different wavelengths are perceived simultaneously as a unified colour. A highly saturated colour consists of a very narrow band of wavelengths, for example, a true red colour, while a less saturated one consist of a wider band of wavelengths, for example, a bluish-red colour [Mayhew, 1992].

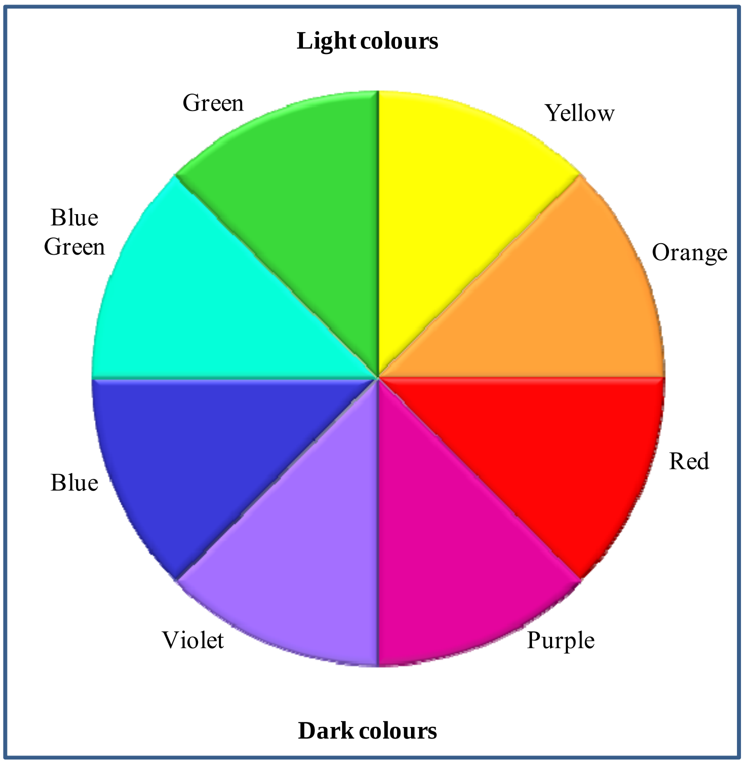


Figure 5.1: Colour hue circle [Arditi, 1999]

3. *Lightness* is a measure of how much light is reflected from a surface relative to the light from nearby surfaces. Colour lightness consists of a series of achromatic or hue-less colours, ranging white through gray to black. An achromatic colour results when all wavelengths are equally reflected (white colour) or equally absorbed (black colour). When a specific hue is combined with achromatic colours, the lightness of the resulting colour depends on the lightness of the achromatic colour [Kotzé and Johnson, 2004; Mayhew, 1992].

Colour blind people have limitation in their ability to distinguish between different colours, especially red and green colours [Dix et al., 2004; Rigden, 1999]. The human eye is composed of a number of structures, including the light-sensitive retina. The retina has two photoreceptor cells, the *rods* and *cones*. Rods are highly sensitive to light and enable us to see when illumination is low, but they cannot distinguish between different colours. Colour is perceived through cone cells, which are less sensitive to light [Dix et al., 2004; Rigden, 1999]. There are three types of cones and each is sensitive to different light wavelengths: the ρ (rho) or 'red'; γ (gamma) or 'green'; and, β (beta) or 'blue' pigments, thus forming the basis for the classification of colour blindness [Rigden, 1999].

In a person with *anomalous trichromacy*, one of the three cone pigmentations has been distorted, thereby limiting three-dimensional or trichromatic vision to some degree. *Dichromacy* is a condition where one of the three cone pigments, usually 'red' or 'green', is missing, limiting vision to two dimensions. One-quarter of colour blind people are dichromats [Rigden, 1999]. In a person with monochromacy, only one of the cone pigments is present. *Trichromats* have all three cone pigments functioning and are able to perceive different colours.
The relationship between trichromatic, dichromatic and monochromatic visions is depicted in Figure 5.2, while Figure 5.3 shows how traffic lights are perceived by normal vision and colour-blind individual who are unable to differentiate between the red and green lights.

Trichromatic (normal) vision | Dichromatic vision | Monochromatic vision

Figure 5.2: The relationship between trichromats, dichromats and monochromats [http://anthro.palomar.edu/primate/color.htm]

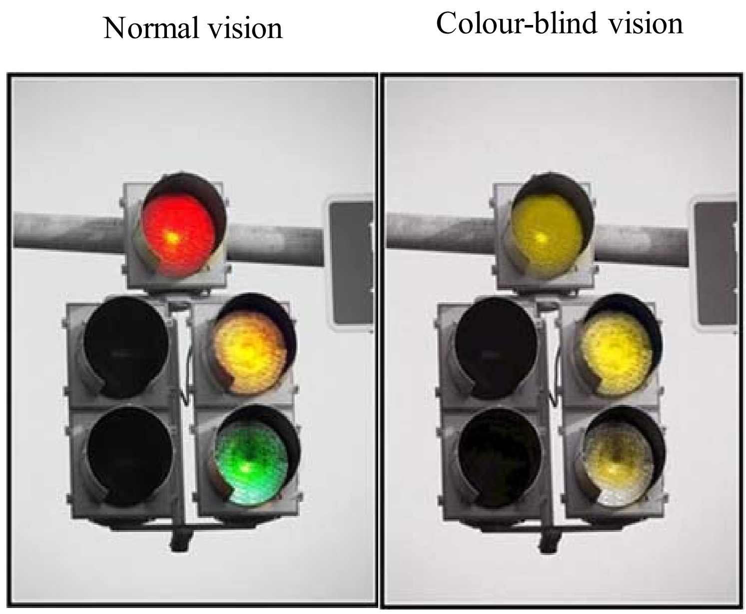


Figure 5.3: Vision by normal and colour-blind persons [http://critiquewall.com/2007/12/10/blindness]

Colour blindness has implication for the design of user-interfaces. A person's ability to effectively distinguish between colour combinations is influenced by partial sight, the effect of aging and congenital colour deficits [Arditi, 1999]. Colour-coding is typically used on the interface to provide visual clue to users by applying colours to words, buttons or screen regions to differentiate functions or to group similar items. People with colour blindness find it difficult to distinguish between shades of red and green colours; hence, the exclusive use of colour for cueing should be avoided [Hoffman, 1999].

Another problem that arises from the use of colour on the interface relates to background and foreground colours. Colour-blind individuals have problem distinguishing between background and foreground colours of low contrast. Arditi [1999] recommends contrasting dark colours from the bottom half of the hue circle (Figure 5.1) against light ones from the top half of the circle, and to avoid contrasting light colours from the bottom half with darks colours from the top half. An effective and ineffective colour contrasts is illustrated in Figure 5.4.

5.3.1.2 Blindness

According to the World Health Organization (as cited by Jacko, Vitense and Scott [2008]), a person is said to be blind if s/he has visual acuity of worse than 20/400. *Visual acuity*, a measure of visual function, is the smallest object that is perceivable by the eye from a given distance. It is expressed as a fraction, where the numerator represents the distance at which a person can recognize an object and the denominator represents the distance from which a normal eye can recognize the same object [Jacko et al., 2008]. The leading cause of blindness in developing countries is cataract, accounting for nearly fifty percent of blindness. Other causes include age-related degeneration and complication of the medical condition, diabetes [World Health Organization, 2009].

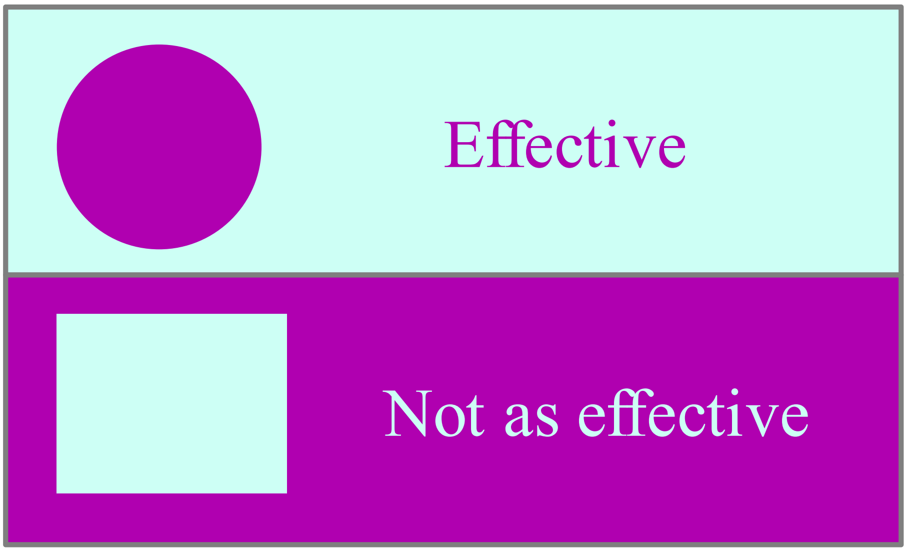


Figure 5.4: Effective and ineffective colour contrast [Arditi, 1999]

Prior to the advent of graphical user interfaces (GUIs), blind users interacted with computer systems using assistive technologies, such as, screen readers, a tool that transforms screen information into synthetic speech, and Braille output displays, a rectangular ruler with pins, capable of protruding and retracting to form Braille characters [Dix et al., 2004; Evers and Hillen, 2007]. The move towards GUIs and the explosive growth of the Internet created accessibility barriers for blind users because screen readers and Braille outputs are restricted in their ability to interpret graphical interface elements [Dix et al., 2004].

Accessibility problems relating to Web site navigation are [Evers and Hillen, 2007]:

- The absence or incorrect use of meaningful textual alternatives for content, such as embedded objects, makes visual elements invisible to assistive technology. When such elements are central to successful navigation of the page, a screen reader will not be aware of their function.
- Inaccessible navigational content, such as, menus and hyperlinks, which cannot be accessed by assistive technology.
- The use of mark-up and layout, such as, colour and size for headings, to convey important information, rather than tags or tables, to structure information leading to information loss.
- Content that require interaction, such as, applets, flash animations and movie instances, have no textual or audio alternatives.

Other problems encountered by blind users while using the Internet include excessive strain on cognitive resources due to large amount of information converted to speech. It is difficult for a blind user to get the overview of an entire Web page because screen information is rendered in a linear fashion. Searching for a link is particularly difficult since all the text on the screen must be scanned by checking each word for its special attributes to determine whether it is a link or not. Returning to a previously located link is equally difficult, as the entire text must be searched again. This process is time consuming and exacerbates the problem of getting lost in cyberspace [Kennel, Perrochon and Darvishi, 1996].

Access to electronic information by blind users can be supported through the use of speech recognition, speech synthesis, and the tactile output [Dix et al., 2004]. One of the principles of universal design, discussed in section 5.5.1.1, is the provision of information using multiple modes such as, text, graphics, sound and touch. The following paragraphs give overview of techniques that could be used to enable access to electronic information by blind users.

- *Speech recognition* is an input mechanism that recognizes human speech and translates it to commands executable by the computer and can enhance human-machine interaction. Although a number of commercial speech recognition systems are available, for example, Dragon Naturally Speaking by Dragon Systems Inc., the success of these systems are still limited due to (i) the complexity of human language, (ii) interference from background noise, and (iii) variation in human speech [Dix et al., 2004; Jacko et al., 2008].

  Despite their limitations, speech as an input method is beneficial to visually impaired users and people without visual disabilities. Speech can be used as an input method when the user's hands are occupied in other tasks, for example in a factory. It is also beneficial to users with physical impairments (see section 5.3.3), which restricts the use of hands [Dix et al., 2004; Mayhew, 1992].

- *Synthesized speech* is "speech that is generated by concatenating basic speech sounds by rule" Simpson et al. [1985] (cited in Mayhew [1992:419]). Synthesized speech can be used as an output mechanism for blind users. Screen readers are implemented using synthesized speech.

  Screen readers have a number of limitations, including the following:

  - They are only capable of reading textual information. When graphical elements are encountered the associated alternative text (*alt text*) is read out. Graphics without *alt text* are not visible to screen readers.
  - Many screen readers read from left to right, beginning at the top of the screen. This means that information that is arranged in columns can become muddled and meaningless.
  - In Web-based applications, it may be difficult for screen readers to follow links that are embedded in text, especially when there are several links in a block of text.

  These limitations have to be taken into account when designing applications that can be accessible to screen readers [Dix et al., 2004; Jacko et al., 2008].

  A major problem of synthesized speech relates to the transient nature of spoken words, making retrieval impossible. The user has to be able to recall previous information. This places a high requirement on users' cognitive resources [Dix et al., 2004; Evers and Hillen, 2007]. Another problem involves the imperfections in speech synthesis and their intrusive nature, necessitating the use of headphones [Dix et al., 2004; Mayhew, 1992].

- *Tactile display* conveys navigation and graphical elements information using electronic Braille display for the Braille literate user. Braille is not constrained by language rules, making it possible for the user to 'read' another language without additional hardware or software [Jacko et al., 2008; Pierce, 2004].

5.3.2 Auditory (Hearing) Impairments

A person is said to be deaf if s/he has substantial hearing impairments in both ears [World Wide Web Consortium, 2005]. A hard of hearing individual has mild to moderate hearing impairment.

Compared to users with visual disabilities, computer-based applications have benefited users with hearing impairments to a greater extent since the visual channel is utilized to convey information. However, the trend towards the use of multi-media presentations with auditory narrations, and increase in the use of sound in interactive systems can lead to accessibility problems for users who are deaf or hard of hearing [Dix et al., 2004; Hanson, 2008].

Contemporary computer applications now combine text, sound, pictures, animations and video to present information. When information conveyed from each of these channels differ from one another, it becomes inaccessible to the hearing impaired user [Hanson, 2008].

5.3.2.1 Techniques to Support Users with Auditory Impairments

- *Captioning* provides textual alternative to speech and sound events. Similar to the sub-titling used in television programs, captioning differs in that it is specifically designed for people with hearing impairments and provides additional comments on the sound that it represents. For example, in addition to the textual rendition of music, a caption will also include a comment like 'music playing' [Hanson, 2008].

  To enable readability, a caption should be synchronized with the audio output and the text should be of correct font and type [David, 2008; World Wide Web Consortium, 1999]. Another issue relating to the design of caption relates to whether to render a verbatim transcript of the audio or provide a simplified version but most people with hearing impairments prefer word for word transmission of the audio output [National Institute on Deafness and Other Communication Disorders, 2002]. In addition to benefiting the hearing impaired, captioning also improves the reading skills of children and can enhance learning a second language [Hanson, 2008].

- *Signing* is a mechanism for representing deaf sign language for computer storage and display to enable input, retrieval and manipulation by people with hearing impairments [Frishberg, Corazza, Day, Wilcox and Schulmeister, 1993]. In addition to access to audio information for deaf and hard of hearing users, signing interfaces are also used to teach reading and writing skills to these users and can be used by people who want to learn sign language.

  Sign language is typically used by some members of the deaf and hard of hearing community in a particular country. The sign language used varies from one country to another and within each country from region to region. In the United State, the language of the deaf community is known as American Sign Language (ASL) while in South Africa it is known as the South African Sign Language (SASL) [Hanson, 2008; National Institute for The Deaf, 2009].

Sign language interfaces can be implemented through a number of techniques. These includes the use of fingerspelling, live signing videos, and signing avatars [Hanson, 2008].

- In fingerspelling, each letter of the alphabet is represented by a unique hand shape. The words to be conveyed are then spelt out letter-by-letter on the hands [Hanson, 2008].
- In live signing, a pre-recorded or live signing video of interface elements is transmitted over high-speed networks. Live signing video is not suitable when the sign version of interface elements must be created in real time [Hanson, 2008].
- A signing avatar is a technique that employs virtual reality to produce computer generated sign language. Signing avatars vary, from those that display only hand movements, to those that are capable of displaying the full signer with facial expressions to convey deeper meanings and emotions [Hanson, 2008; Sansonnet, Braffort, Martin and Verrecchia, 2009].

  Unlike signing videos, which rely on pre-recorded materials, signing avatars are capable of generating new materials without requiring new recordings. Also, they do not require high bandwidth for transmission and can be viewed from any angle by the user [Hanson, 2008; Verlinden, Tijsseling and Frowein, 2002].

5.3.3 Physical Impairments

Physical impairments affect a user’s ability to interact physically with computing technologies. Contemporary computing devices are becoming smaller, with the potential for creating accessibility barriers for users with physical impairments, either due to their inability to operate the device or control it with the necessary precisions [Dix et al., 2004; Sears, Young and Feng, 2008].

Physical impairments can be caused by congenital malformations, for example, congenital absence of an arm or a leg (Amelia), or health related conditions, such as, spinal cord injuries, arthritis and cerebral palsy. Users with these types of physical impairments find it difficult to use standard computer keyboard and mouse [Feng and Sears, 2007; Sears et al., 2008].

There are four categories of physical impairments affect the use of computing technologies [Sears et al., 2008]:

1. *Structural deviations* result from malformation or loss of body parts, for example the absence of a finger, hand or an arm, or due to deviation from the normal positioning of a body part.
2. *Mobility functions* affect a person’s ability to move a joint or bone; thus limiting the range of possible movements and the ease with which such movements can be made. For example, a person with arthritis may find it difficult to bend the fingers.
3. *Muscle power functions* affect a person’s ability to generate necessary force required to contract a muscle or a group of muscles. Muscle power loss is caused by brain or spinal cord injuries. Loss of muscle power can be partial (paresis) or complete (paralysis).

4. *Movement function* affects a person's ability to control voluntary and involuntary movements.

5.3.3.1 Interaction Techniques to Support Users with Physical Impairments

- *Access through modification to standard keyboard:* Modifications to standard keyboard is sometimes all that is required to provide accessibility to users who can still use their hands but lack the precision required to use the mouse. Using keyboard accelerators, the user can access system features and functionalities by pressing different key combinations. This can be difficult, and sometimes impossible, for a user with the use of only one hand. For example, locking the workstation requires the user to press the *Alt* and *Delete* keys while holding down the *Ctrl* key on the keyboard.

  Contemporary operating systems, for example the Windows operating system, provide features that facilitate access to users with physical impairments. Common built-in keyboard accessibility features are [Bergman and Johnson, 1995; Kotzé, Eloff, Adesina-Ojo and Eloff, 2004; Vanderheiden, 1994]:

  - *StickyKeys* allow modifier keys, such as, *Shift* and *Alt* to remain active while the user presses other keys one at a time.
  - *RepeatKeys* enable the user to adjust the rate at which key presses are displayed on the screen, thus eliminating unwanted multiple characters.
  - *SlowKeys* prevent accidental keypresses by allowing the user to set a delay time for the keyboard so that key presses are only accepted when the specified time has elapsed.
  - *BounceKeys* prevent the keyboard from accepting quick consecutive press of the same key, which typical occur in users with tremors.
  - *MouseKeys* allow the user to control cursor movements and mouse button functions, such as a click effect, using the keyboard.
  - *ToggleKeys* provide a sound which alerts the user when locking keys, such as the *Caps Lock*, *Num Lock* or *Scroll Lock* key, is pressed.

  Many users with physical impairments are not aware of these built-in keyboard accessibility features and are thus unable to benefit from them. Kotzé et al. [2004] found that only 56 percent of participants in a study on the use of computer systems by quadriplegics knew about StickyKeys or other accessibility features in the Windows operating system.

- *Head-controlled interaction:* Head-controlled interaction devices enable users with limited or no use of their hands to control cursor movements and generate text using their heads [Feng and Sears, 2007]. Head pointing devices, based on the standard mouse, have to be physically mounted on the head and depend on the ability of the user to move the head with precision and must have adequate neck control [Mauri, Granollers, Lorés and García, 2006]. However, text generation is very slow with head-controlled device and it may cause neck muscle fatigue. In addition, to be effective, the device must be calibrated [Feng and Sears, 2007].

- *Eye-controlled interaction:* Eyegaze systems enable people with limited or no control of their hands to interact with a computing system just by looking at it. These systems vary from those that require the user to wear specialized glasses or a head-mounted device to those built into the screen and less invasive. An eyegaze system facilitates interaction by beaming a low-power laser into the eye, which is then reflected off the retina. Eyegaze systems must also be calibrated and they are less precise for tasks requiring fine movements, for example, drawing.
- *Speech-based interaction:* Speech recognition systems enable users with physical impairments to interact with computing systems by issuing commands through spoken words, which are executed by the computer (see section 5.3.1.2.1). Speech can be used as a navigation mechanism to access system resources and for text generation.
- *Onscreen keyboard:* A virtual representation of the standard or modified keyboard on the screen for users with physical impairments unable to type with the standard keyboard. Keys on onscreen keyboard can be accessed though the mouse, joystick or other pointing devices, such as a head pointer [Kotzé et al., 2004; Mauri et al., 2006].

Standard onscreen keyboards, for example Windows onscreen keyboard (Figure 5.5), are based on the assumption that the user is able to use pointing devices with sufficient precision. However, it is often the case that users with physical impairments like cerebral palsy and spinal cord injuries also have difficulties with fine motor coordination. These users may find it difficult to position the pointer over onscreen keys and even after positioning the pointer it can be difficult to click and type or keep the pointer still without moving it to another area of the screen. This can lead to inadvertent keystrokes and can slow down the typing speed.

To address some of the problems associated with standard onscreen keyboard, Kjeldsen [2007], developed an onscreen keyboard (Figure 5.6) specifically for users with poor pointer control. Rather than follow the conventional alpha-numeric keyboard layout, the keys are displayed along the top edge of the screen whenever the pointer is positioned in this region. ScreenEdge allows users to select keys by moving the pointer from left to right, and to insert the selected keys by moving the pointer in a downward position. ScreenEdge offers a number of advantages over standard onscreen keyboard. It is more effective because the user can force the pointer against the edge of the screen; key selection is also easier.

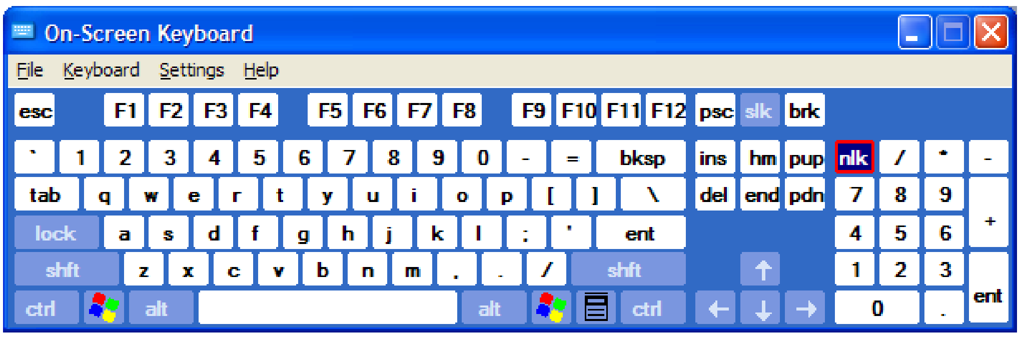


Figure 5.5: Windows onscreen keyboard

Figure 5.6: ScreenEdge keyboard [Kjeldsen, 2007]

5.3.4 Cognitive Impairments

Cognitive impairments is the general term used to describe a variety of conditions, including those affecting thinking, memory, learning and perception [Vanderheiden, 1994]. Cognitive impairments could occur as a result of birth defects or congenital malformations, following injuries, stroke, and the effect of aging.

Of the four categories of impairments (visual, auditory, physical and cognitive) and their implications for the design of interactive systems, cognitive impairments have received the least attention from HCI practitioners for the following reasons [Keates, Kozloski and Varker, 2009]:

- Cognitive impairments are invisible compared to other categories of impairments. It is easier to identify someone with paralysis of upper and lower limbs, for instance, than it is to identify someone with dyslexia.
- Cognitive impairments are difficult to diagnose. Although some symptoms of cognitive impairments can be identified, for example through memory tests, others cannot be diagnosed.
- There is no universal agreement on the definition of, and what should be categorized as cognitive impairment. There are researchers who believe that learning difficulties and behavioural disorders should be treated separately from cognitive impairments.
- Perhaps the most important reason is the fact that it is difficult to design applications to support users with cognitive impairments due to lack of understanding of the problem area to be addressed.

Despite the problems associated with the design of applications to support users with cognitive impairments, ICTs have great untapped potential to enhance the quality of life of people with cognitive impairments, ICTs can enable them to (i) retain a high level of independence and control over their lives, (ii) provide appropriate levels of monitoring and supervision of high-risk individuals (for example dementia) without violating their privacy, (iii) ability to keep them both physically and intellectually, and (iv) enable communication and social interaction to reduce social exclusion [Newell, Carmichael, Gregor, Alm and Waller, 2008].

5.3.4.1 Strategies to Support Users with Cognitive Impairments

Design principles to guide the design of applications to support users with cognitive impairments include:

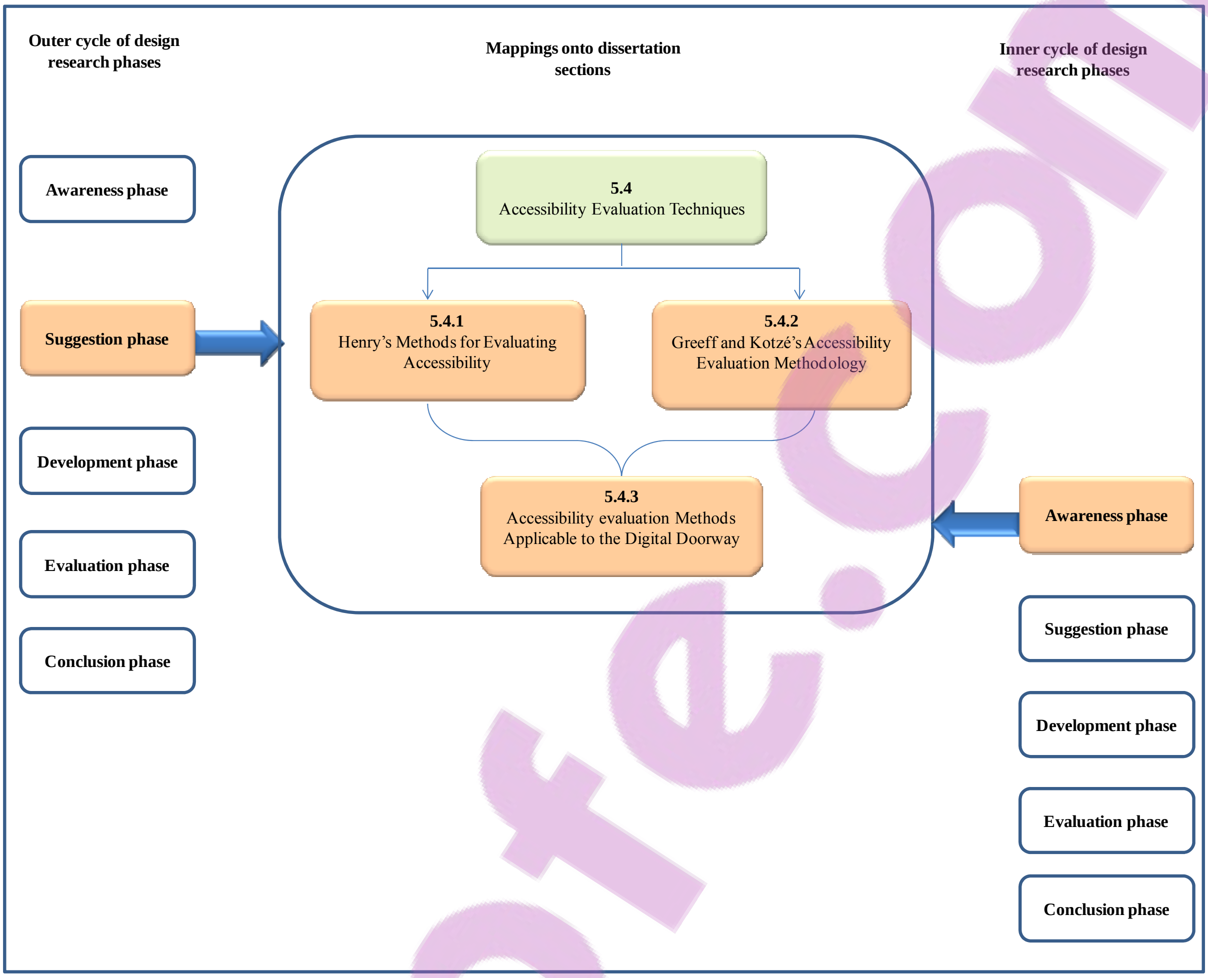
- Users with cognitive impairments should be allowed to complete every aspect of a task at their own pace because they tend to be slower when compared with other users. This can boost the morale of the cognitive impaired user who has constantly been told to 'hurry up'. The computer can thus be seen as being more patient [Newell et al., 2008].
- Interface elements should be displayed using simple, clear and unambiguous terms. There should be consistency in the location and layout of interface elements to enable easy recognition and access [Keates et al., 2009; Newell et al., 2008; Vanderheiden, 1994].
- The grouping of elements, for example menus, should be kept to minimum. Users should be presented with only a single item at any point in time. The presentation of multiple information in parallel should be avoided as this increases the possibility of user error [Keates et al., 2009; Newell et al., 2008].
- The use of a multimodal technique to present information using graphics and sound in addition to text is beneficial to users with cognitive impairments. Whenever possible, the graphics should be concrete representation of elements being presented to aid recognition [Newell et al., 2008].

Having discussed the various forms of disabilities and the ways in which the design of interactive systems can affect people with disability, section 5.4 will examine the various techniques that could be used to evaluate interactive systems' accessibility to determine the appropriate methods that can be utilized in the evaluation of the applications and interfaces on the DD.

5.4 ACCESSIBILITY EVALUATION TECHNIQUES

This section maps onto the suggestion phase of the outer cycle of the design research and the awareness phase of the inner cycle. The suggestion phase examines accessibility evaluation methods proposed by Henry [2007] and the Web accessibility methodology by Greeff and Kotzé [2009], to determine the appropriate method for evaluating direct accessibility support provided in the DD while the awareness phase reveals the need for heuristics that address direct accessibility.

Mapping of section 5.4 to the design research phases



5.4.1 Henry's Methods for Evaluating Accessibility

The important role of evaluation in the development of interactive systems cannot be overemphasized. Accessibility evaluation can be incorporated into some of the usability evaluation methods discussed in 4.2.3. This section examines six methods that could be utilized in order to evaluate the accessibility of interactive system.

1. *Standards Review*: The target application or system is assessed for conformance to specific design standards, which can be internally developed within an organization or externally specified by a national or international organization, for example the WCAG 1.0. [1999]. Accessibility standards review is more thorough when it concerns conformance to legal requirements [Henry, 2007].
2. *Heuristic Evaluation*: Involves expert evaluators examining an application's interface to independently judge its compliance with a set of evaluation heuristics (see section 4.2.3.1). In accessibility evaluation, the heuristics address accessibility characteristics [Henry, 2007].

3. *Design Walkthroughs:* Evaluators mimic the activity of a representative user while being guided by a design team member. The method is used early in the design process to uncover potential accessibility problem.

   Design walkthroughs can be conducted for accessibility evaluation either by focusing on specific accessibility concerns as part of regular development walkthroughs (for example checking that a task requiring mouse clicking action by the user can also be accomplished through the keyboard) or by conducting the walkthrough specifically for accessibility evaluation.

   Conducting design walkthroughs specifically for accessibility involves the use of personas with disabilities and scenarios that require the use of adaptive technologies to complete a task. A persona is hypothetical archetypes of real users that enable designers envision the characteristics of target users for a specific application and has real name, demographic and prior experience information, and the type of disability. A scenario provides the description of how a persona with disability completes a task using the target application with some form of adaptive strategy or an assistive technology, for example, a screen reader [Henry, 2007].

4. *Screening Techniques*: Evaluators interact with a design by removing or modifying one or more physical or sensory abilities. For example, an evaluator might put on a thick glove to reduce hand dexterity or wear a low vision glass to limit vision. It may also involve the use of adaptive strategies or assistive devices. Screening technique is used early in design to reveal potential accessibility problems in a design when the cost and efforts required to make corrections are less than when they are discovered at a later stage. In addition, the method makes later testing with real users with disabilities more efficient by eliminating initial problems before formal testing to ensure judicious use of participants with disabilities, who are sometimes difficult to recruit.

   One limitation of screening technique is the potential to produce incomplete or inadequate results. For example, an evaluator with limited expertise in the use of screen readers might conclude that a given problem is related to the design being evaluated when the problem is due to his/her lack of skills in using the device [Henry, 2007].

5. *Usability Testing*: Involves modifications to the standard usability testing through the involvement of users with disabilities as test participants. The number and characteristics of participants that are included in the evaluation varies, depending on the disability category the application is targeted. However, variation also exists within any specific disability category. For example, an application may be accessible to a blind user who uses a screen reader but not accessible to a user with colour blindness; yet these two users are both classified as having visual impairments [Henry, 2007].

6. *Use of Accessibility Evaluation Tools*: Used to evaluate Web pages and elements of software for compliance to accessibility guidelines and standards, for example the WCAG1.0 and Section 508. Tools, such as Total Validator [2009], can automate initial evaluation but human evaluation through heuristic evaluation, for example, is still required. A software tool can indicate a missing

alt text for an image but it cannot determine whether an existing *alt text* conveys the same information as the image it represents [Henry, 2007].

5.4.2 Greeff and Kotzé's Accessibility Evaluation Methodology

Greeff and Kotzé [2009] propose a lightweight methodology for evaluating the accessibility of interactive systems, including Web-based applications. The methodology involves three iterative phases: an initial evaluation, to determine the level of conformance with established accessibility guidelines such as the WCAG1.0; testing with real users; and, development of context-specific, in-house guidelines for future development within the organization. The three phases are [Greeff and Kotzé, 2009]:

1. *Initial Accessibility Evaluation*: This phase aims to discover accessibility problems inherent in the target application before involving users in evaluation, using automatic software tools to assess the following accessibility issues:
   - Checking whether the Web site conforms to specific accessibility guidelines.
   - Checking the readability of text on the site to ensure they can be understood by target audience of the Web site.
   - Checking the pages' colour contrasts to ensure that the site is usable to people with visual impairments.
   - Checking whether the Web site is accessible to users with visual impairments who make use of screen readers.

   After the initial automated evaluation, identified accessibility problems are corrected. This is then followed by another iteration of the automated evaluation procedure before proceeding to the second phase.
2. *Testing with Real Users*: Evaluation with automatic software tools has limitations (see section 5.4.1). An automatic evaluation tool is as good as the guidelines it is based on. To ensure accessibility and user acceptance by the target user group, the initial automatic assessment with software tools is followed with user testing, where users from different disability categories are involved in the evaluation.
3. *Development of In-House Context-Specific Guidelines*: Involves the development of in-house, context-specific guidelines that can be used by non-expert developers to guide future enhancement and development efforts so that similar design errors are not repeated. These guidelines are influenced by the lessons learnt from the first and second phases of evaluation.

5.4.3 Accessibility Evaluation Methods Applicable to the Digital Doorway

Within the context of this study, the standards review method, proposed by Henry [2007] cannot be used to evaluate the DD, since the project team do not currently use any existing accessibility standards against which it can be assessed.

Design walkthroughs and screening techniques [Henry, 2007] are used early in the design process for formative evaluation. This study is a summative evaluation of an existing, fully functional system; hence these two methods will not be appropriate.

Usability testing with disabled users cannot be performed since the DD does not support indirect accessibility through the use of assistive technologies, making the inclusion of blind users as study participants inappropriate.

The use of automatic software tools will also not be appropriate for evaluating the DD since these tools are typically used for Web-based applications running on Windows operating system. Examples of such tools includes *Cynthia Says* [HiSoftware, 2009] and *Web Accessibility Inspector* [Fujitsu, 2009]. Applications on the DD run on the Ubuntu Linux operating system. These limitations make the use of automatic software tools inappropriate for evaluating the DD.

The first phase of Greeff and Kotzé's [2009] methodology also involves the use of automated software tools, the use of which has been eliminated as inappropriate for evaluating the DD.

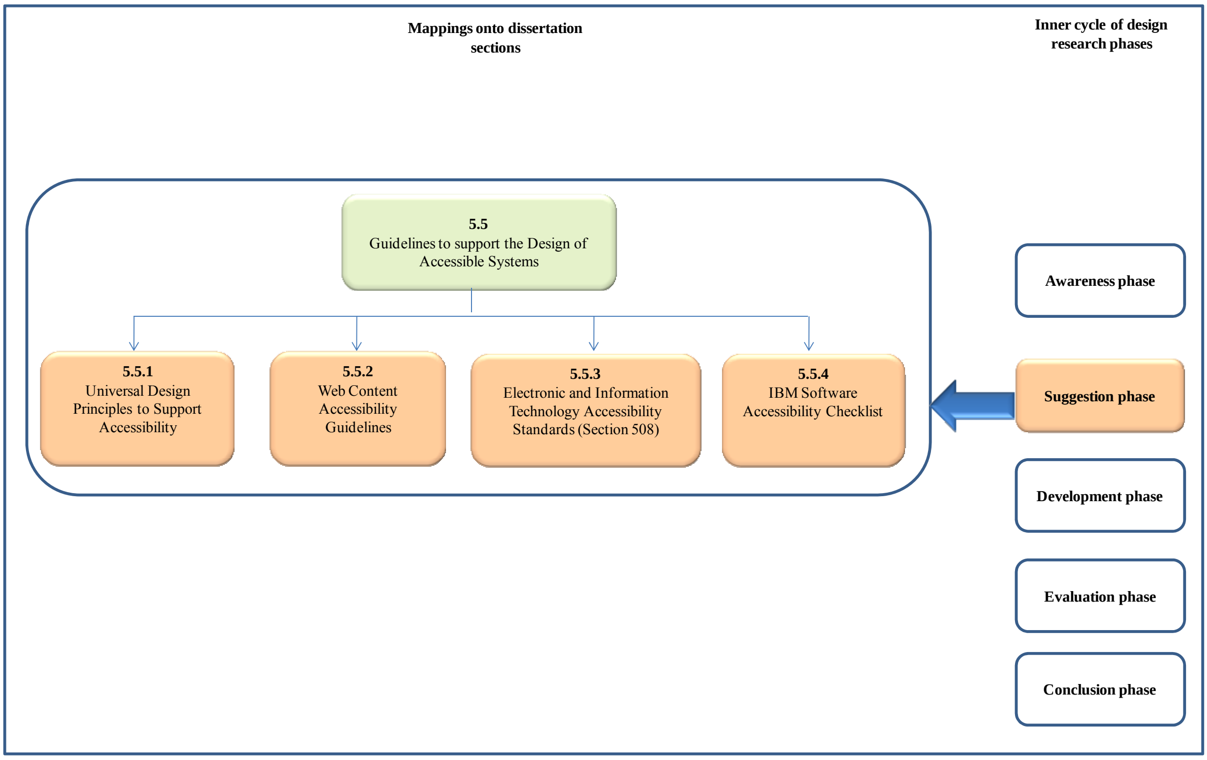
An applicable evaluation method is the heuristic evaluation method. A large part of this study involved literature investigation of existing principles and guidelines for the design of usable and accessible interactive systems. The accessibility design guidelines discussed in section 5.5 will guide the development of accessibility-specific heuristics to determine the level of direct accessibility support provided in the DD.

The third phase of Greeff and Kotzé's [2009] methodology is also relevant; the evaluation heuristics developed in this study will be a useful tool to guide future development of applications for the DD.

5.5 GUIDELINES TO SUPPORT THE DESIGN OF ACCESSIBLE SYSTEMS

In section 5.4, I examined the methods that could be used to evaluate the accessibility of interactive systems. The heuristic evaluation method was identified an appropriate method to evaluate direct accessibility support provided in the DD. The use of heuristic evaluation method requires specific heuristics that focus on accessibility, hence the need to review existing principles and guidelines on accessibility to determine their applicability to the DD context. This section maps onto the suggestion phase of the inner cycle of the design research and will examine the universal design principles of Story, Mueller, and Mace [1998], the Web Content Accessibility Guidelines (WCAG 1.0) developed by the World Wide Web Consortium (W3C) [1999], the United States' electronic and information technology accessibility standards [2000], and IBM software accessibility checklists [IBM, 2009]. The guidelines that are found to be applicable will be highlighted at the end of the sections where they are discussed.

Mapping of section 5.5 to the design research phases



5.5.1 Universal Design Principles to Support Accessibility

Universal design or design for all, is "the process of designing products so that they can be used by as many people as possible in as many situations as possible" [Dix et al., 2004:366]. Diversity in terms of age, abilities, cultural and educational backgrounds, and experiences affect the way people interact with computer systems.

There are two types strategies for the design of universal interfaces, based on the usage of the specific application [Lazar, 2007]. Some applications are targeted at a variety of users regardless of their age, experience or ability. Examples of such applications include government and news Web sites, educational institution Web sites and online communities. Other applications are aimed at specific user groups with special needs, such as, people with Alzheimer's disease, children with autism and users with spinal cord injuries. Universal design aspires to limit the use of multiple interfaces for applications that are targeted at diverse user population.

Universal design is not a 'one size fits all' solution, neither is it design for a niche market of the disabled. Rather, it is the design of applications that (i) could be used by most potential users without any modification, (ii) are easily adaptable to different users through customization, and (iii) has standardized interfaces so that they are accessible through assistive technologies [Darzentas and Miesenberger, 2005]. Universal design can be graphically illustrated through the usability pyramid Nordby [2003] (as cited by Darzentas and Miesenberger [2005]).

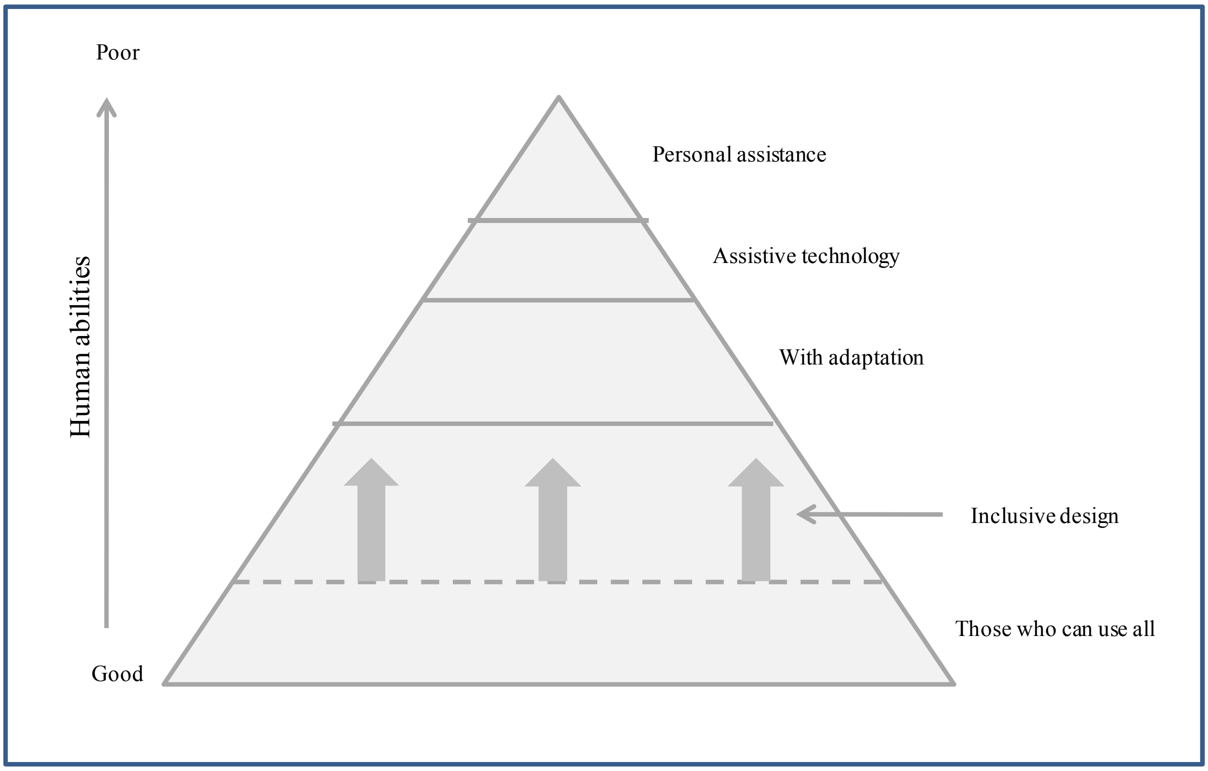


Figure 5.7: Nordby's usability pyramid (as depicted by Darzentas and Miesenberger [2005])

The usability pyramid (Figure 5.7) represents users of interactive applications. Human abilities are shown along the vertical axis with good ability at the bottom and the poorest at the topmost end. The pyramid is divided into four segments. At the base are population majority, who are able to access all applications and devices directly. Immediately above this population segment is a smaller group of people, who can only access applications and devices through some form of adaptation, for example, a magnifier for reading screen text. At the next level is the user group who can only access applications and devices through some form of assistive technologies, such as a screen reader. The peak of the pyramid represents people who can only access applications and devices through the assistance of a third party.

The primary goal of universal design is to shift the boundary between those who can use all applications and devices, and those who can only use the devices with adaptation as far up as possible. This is shown by the upward arrows between the two user groups.

5.5.1.1 Universal Design Principles by Story, Mueller, and Mace

Story, Mueller, and Mace [1998], researchers at the North Carolina State University in the United State developed seven universal design principles which provide a framework for the design of applications that are accessible to many people in many situations. The principles can be widely applied to all areas of design, including interactive computing applications:

1. *Equitable use*: Applications should be designed such that they are usable by people of diverse abilities. No user should be excluded or stigmatized. Whenever possible, access should be identical for all users but when this is not possible, equitable use should be supported. Support for privacy, security and safety should be provided for all users.
2. *Flexibility in use:* The design should support different abilities and preferences by adapting to the user's pace and precision.
3. *Simple and intuitive use:* The system should be easy to understand, regardless of user's knowledge, experience, language or level of concentration. The design should be simple, support users' literacy and language skills and meet their expectations. The system should be organized so that it enables easy access to important areas and provide effective prompts and feedback during and after task execution.
4. *Perceptible information:* The system should communicate information effectively to the user irrespective of the prevailing environmental conditions or the user's ability. The system should provide information using multiple modes, such as graphics, text, sound and tactile modes. Important information should be easy to distinguish from peripheral contents. The system should support the wide range of devices and techniques used to access information by people with different sensory abilities.
5. *Tolerance for error:* The design should lessen the adverse effects that could occur from users' mistakes or unintended behaviours. Actions that could potentially result in dangerous consequences should be hard to initiate. Constraints should be used at strategic points as error prevention mechanisms. Warnings should be provided for hazardous situations.
6. *Low physical effort:* The system should be designed such that it is comfortable to use, minimize physical effort and fatigue. The design should enable users to maintain their natural posture while interacting with the system. Repetitive or sustained physical actions should be kept to minimum.
7. *Size and space for approach and use:* The placement of the system should enable easy reach by any user regardless of body size, posture or mobility. Essential elements and information should be in the line of sight for both standing and seated users. All components should be easy to reach by standing and seated users. The system should allow for variation in hand size and provide adequate space for the use of assistive devices.

5.5.1.2 Applicability of the Universal Design Principles to the Digital Doorway

Five of the seven universal design principles by Story et al. [1998] can guide the development of heuristics for evaluating the direct accessibility support provided by the DD. One principle that is not relevant to the DD context is principle two, *flexibility in use*. This principle relates to the ability of a system to support users with different abilities, especially users with different types of physical disabilities. As stated in section 1.2, the DD does not currently support indirect accessibility through the use of assistive devices or system modifications.

Principle seven, *size and space for approach and use*, is concerned with ergonomic considerations, which is beyond the scope of this study. However, a wheel chair-accessible version of DD is available on request.

Relevant principles include those relating to perceptible information and low physical effort. The provision of perceptible information through the use of multiple modes will ensure that more diverse users are able to access the DD. Universal design principles that are applicable to DD context are summarized in Figure 5.8.

1. *Equitable use:* Applications should be accessible to users regardless of their environment of usage.
2. *Simple and intuitive use:* The level of the system's complexity should be low, taking into account users' age and experience.
3. *Perceptible information:* Information should be provided through multiple modes to ensure access and comprehension.
4. *Tolerance for error:* The system should prevent unintended user errors through strategic use of constraints.
5. *Low physical effort:* Users should be able to access the system without undue physical efforts.

Figure 5.8: Universal design principles applicable to Digital Doorway context

5.5.2 Web Content Accessibility Guidelines

The W3C is an international organization with members from corporations, academia and research institutes. It develops standards that will enable the realization of the full potential of the Web. The Web Accessibility Initiative (WAI) is an affiliate of the W3C and aims to improve Web accessibility to people with disabilities [World Wide Web Consortium, 2009]. Web Content Accessibility Guidelines (WCAG) are developed by the WAI to guide the development of accessible Web sites.

The first version of the guidelines, WCAG 1.0 [1999], has fourteen guidelines, each guideline has a number of checkpoints describing its in typical content development scenarios. Each checkpoint has a priority level assigned to it, depending on its impact on accessibility. A checkpoint with Level 1 priority is the most basic requirement which must be satisfied, otherwise some user groups will not be able to access the Web page. A Level 2 priority checkpoint should be satisfied, else some user groups will find it difficult to access the information. Level 3 priority checkpoints may be satisfied, otherwise some user groups will find it somewhat difficult to access information in the Web page.

The priority levels are used to determine the degree to which a given Web site conforms to the WCAG. A Web page that satisfied all Level 1 priority checkpoints is assigned the conformance level 'A'. When all the Level 1 and Level 2 priority checkpoints have been satisfied, the conformance level 'AA' is assigned while adherence to all three priority level checkpoints will earn a Web page the conformance level 'AAA'.

A second version of the guidelines, WCAG 2.0 [2008], has been developed. It also provides recommendations for the development of an accessible Web content. The guidelines consist of four principles: perceivable; operable; understandable; and, robust. Each principle has a number of specific sub-principles.

I decided to utilize WCAG 1.0 [1999] to provide guidance for the development of direct-accessibility evaluation heuristics for the DD for two reasons:

1. The second version of the guidelines, WCAG 2.0, was still under development during the development phase of evaluation heuristics for the DD (see section 3.4.1.2). Version one was a more stable version.
2. WCAG 1.0 is the version currently being used for other applications at the Meraka Institute. An example of a Meraka application using the WCAG 1.0 is the National Accessibility Portal (NAP).

The fourteen guidelines contained in the WCAG 1.0 are [World Wide Web Consortium, 1999]:

1. *Provide equivalent alternatives to auditory and visual contents:* All graphical, video and audio information should have textual equivalents.
2. *Do not rely on colour alone:* Colours should be used only as enhancements and not the primary method for displaying important information because colour-blind users and people whose devices are unable to display colour will not be able to access the information.
3. *Use markup and style sheets and do so properly:* Document markups and the control of information presentation using style sheets are important to enable access to users with specialized applications like screen readers.
4. *Clarify natural language usage:* Markups should be used on natural language so that when changes are made, user agents can automatically switch to the new language. Markups also allow search engines to locate the keywords and documents in a desired language.
5. *Create tables that transform gracefully:* Markups should be used on tables to enable access to Web browsers and other user agents.
6. *Make sure that pages featuring new technologies transform gracefully:* Pages that use new technology should also be accessible to people using older technology.
7. *Ensure user control of time-sensitive content changes:* Users should be able to pause or stop any moving, blinking, scrolling, or automatically updated objects. This will allow slow users to read the information at their own pace.
8. *Ensure direct accessibility of embedded user interfaces:* Any embedded object, for example, an applet which has its own interface, should be accessible as the browser itself.
9. *Design for device-independence:* A Web page should support users' interaction through a variety of input and output devices to facilitate access to diverse users.
10. *Use interim solutions:* Interim accessibility solutions should be used to enable access to users with assistive technologies and older browsers.

11. *Use W3C technologies and guidelines:* W3C technologies and accessibility guidelines should be used correctly. Whenever this is not possible, an alternative version of the content should be provided.
12. *Provide context and orientation information:* Web page elements should be grouped logically and contextual information on their relationships should be provided to facilitate easy comprehension, especially for users with cognitive and visual impairments.
13. *Provide clear navigation mechanism:* Navigation tools and orientation information should be provided in clear and consistent manner so that users are can locate the information.
14. *Ensure that documents are clear and simple:* The layout of the Web page should be simple, clear and consistent to enable easy comprehension especially for users with cognitive impairments and reading difficulties.

5.5.2.1 Applicability of the WCAG 1.0 to the Digital Doorway

The WCAG 1.0 [1999] is primarily designed for Web accessibility through the use of assistive devices and system modifications. The majority of these guidelines: guidelines three to six; guideline eight; and, guidelines ten to thirteen are specifically concerned with technical conformance of Web-based applications and are not relevant to the DD context.

Five of the guidelines can be used to develop heuristics for evaluating direct accessibility support provided in the DD. WCAG 1.0 [1999] guidelines that are relevant to the DD context are summarized in Figure 5.9.

1. *Provide equivalent alternatives to auditory and visual contents:* This guideline is interpreted in a similar way to that relating to the provision of information using multiple modes, such as text, graphics and audio.
2. *Do not rely on colour alone:* Colours should not be used as the main method for representing and displaying important information.
3. *Ensure user control of time-sensitive content changes:* Users should have control over the speed at which any moving, blinking and audio information is displayed. They should be able to pause, continue, or repeat these types of information.
4. *Design for device-independence:* Systems should be designed such that it supports user interaction through different input and output devices like the mouse and keyboard, to enable access by diverse users.
5. *Ensure that documents are clear and simple:* The layout of interface elements should be simple, clear and consistent so that users with cognitive impairments can comprehend the information.

Figure 5.9: WCAG 1.0 applicable to Digital Doorway context

5.5.3 Electronic and Information Technology Accessibility Standards (Section 508)

The United States electronic and information technology accessibility standards, (Section 508) is the 1998 amendment to the Rehabilitation Act, which compels all federal agencies to make electronic and information technologies accessible to people with disabilities. The legislation applies to the development, procurement, maintenance or use of ICT infrastructures, technical standards for software applications and operating systems, Web-based Intranet and Internet applications, telecommunication products, video and multi-media products, self-contained closed products, such as information kiosks, and desktop and portable computers. Section 508 aims to remove barriers in access to information by people with disabilities and create new opportunities for them [Waddell, 2002].

There are six categories of technical standards [United States Access Board, 2000]:

1. Software applications and operating systems.
2. Web-based Intranet and Internet information and applications
3. Telecommunications products.
4. Video and multimedia products.
5. Self contained, closed products.
6. Desktop and portable computers.

The software applications and operating systems standards are [United States Access Board, 2000]:

1. System functionalities shall be accessible via the keyboard in a software application that runs on a system that has keyboard, especially when the functionalities or their results can be perceived textually. This standard is similar to guideline nine of WCAG 1.0 (section 5.5.2) and IBM checklist 1.1 (section 5.5.4).
2. Accessibility features built into products that comply with industry standards shall not be disabled by developers. Accessibility features built into operating systems shall not be disabled if the application programming interface (API) for the features has been documented by the manufacturer and is available to developers. This standard is similar to IBM checklist 1.2.
3. A well-defined, on-screen indication of the current focus shall be provided and move among interactive interface elements as the input focus changes. The focus shall be programmatically exposed so that assistive technology can track focus and focus changes.
4. Sufficient information about a user element, such as, the identity, the operation and state of the element shall be available to assistive technology. Images that represent program elements shall also be available in text.

5. When bitmap images are used to identify controls, the status indicators and meaning assigned to the image shall be consistent throughout the application's performance. This standard is similar to the general usability principle of consistency.
6. Textual information shall be provided through operating system functions for displaying text. At least, information on the text content, text input caret location and text attributes shall be provided.
7. Applications shall not override user-selected contrast, colour selections and display attributes.
8. Animated information shall be available in at least one non-animated presentation mode when required by the user.
9. Colour coding shall not be used as the only means to convey information, indicate actions, prompt user response, or distinguish visual elements.
10. Applications that allow users to adjust colour and contrast settings shall also provide a variety of colour selections that can produce a range of contrast levels.
11. Software applications shall not use flashing or blinking text, objects, or other elements that has a flash or blink frequency range of between 2 and 55 Hz.
12. Electronic forms shall allow people using assistive technology to access information, field elements and the functionality required to complete and submit the form.

5.5.3.1 Applicability of Section 508 to the Digital Doorway

Although the name, software applications and operating systems, suggests the standard should be highly applicable to the DD context, only four of the twelve standards in this category were found to be relevant in deriving heuristics to evaluate direct accessibility support provided by the DD.

Section 508 standards that are not applicable to DD context are:

- Standards two, seven, eight, ten, and eleven: These standards are concerned with functionalities that are not provided by the DD.
- Standards four, six and twelve: These standards deal with the provision of indirect accessibility support through the use of assistive devices. As stated in section 1.4, the DD does not support the use of assistive devices.

The relevant part of the software applications and operating systems part of Section 508, which could be used to formulate heuristics for evaluating direct accessibility support in the DD, is summarized in Figure 5.10.

1. Users should be able to access system's resources and provide input through the keyboard.
2. A visible cue should be provided for the current focus of interface elements, for example, form fields.
3. The meanings of similar images, control buttons, and other interface elements should be consistent throughout an application's performance.
4. Information should not be provided through colour coding alone.

Figure 5.10: Section 508 standardards applicable to Digital Doorway context

5.5.4 IBM Software Accessibility Checklist

The IBM accessibility checklists, based on the United State electronic and information technology accessibility standards and W3C guidelines, are designed to help developers understand the need for providing accessibility supports to users with disabilities [IBM, n.d]. The current version of the checklist (Version 3.6) is grouped into five categories. Each category has a number of specific checklists or guidelines and a rationale for the guideline [IBM, 2009].

The five primary checklist categories and their associated guidelines are [IBM, 2009]:

1. Keyboard access
   1.1. *Provide keyboard equivalents for all actions:* this is to enable users who do not have the level of precision required to position the pointer when using a mouse. Keyboard equivalents also allow blind users to navigate interface elements.
   1.2. *Do not interfere with keyboard accessibility features built into the operating system:* so that users can customize the features according to their needs.
2. Object information
   2.1. *Provide a visual focus indicator that moves among interactive objects as the input focus changes. This focus indicator must be programmatically exposed to assistive technology:* so that a screen reader can determine the current location which must be read to the user.
   2.2. *Provide semantic information about user interface objects:* textual equivalents of images should be provided. This guideline is similar to the first guideline of WCAG 1.0.
   2.3. *Associate labels with controls, objects, and images:* interface elements should have meaningful captions. There should be consistency in naming of similar elements throughout the application.
   2.4. Electronic forms should allow people using assistive technology to access information, field elements and the functionality required to complete and submit the form.
3. Sounds and multi-media
   3.1. *Provide an option to display a visual cue for all audio alerts:* this is because audio alerts may not be detected by deaf or hard of hearing users and people with situational limitations.

3.2. *Provide accessible alternatives to significant audio and video:* multi-media applications that utilize audio and video content should provide synchronized alternative formats.

3.3. *Provide an option to enable users to adjust the volume:* users should have control over appropriate levels of audio outputs to suit their requirements and situations.

4. Display

4.1. Provide text through the standard system function calls or an application programming interface which is able to interact with assistive technologies.

4.2. *Use colour as an enhancement, not as the only way to convey information or indicate an action:* because colour-blind users are unable to differentiate between different shades of colour.

4.3. *Support system settings for high contrast for all user interface controls and client area content:* do not override the contrast settings that are provided by operating systems because low-vision users require high contrast between a text and its background to read information.

4.4. *When colour customization is supported, provide a variety of colour selections capable of producing a range of contrast levels:* applications that allow users to customize colours should provide different levels of contrast and colour choices.

4.5. *Inherit system settings for font, size, and colour for all user interface controls:* operating systems and software applications usually allow users to customize font, size and colour settings. Applications should allow users to select their own preferred setting rather than the custom setting provided by the application.

4.6. *Provide an option to display animation in a non-animated presentation mode:* users should be able to disable animations so that those using assistive technology like screen reader can access the information.

5. Timing

5.1. *Provide an option to adjust response times on timed instructions or allow the instructions to persist:* an application that requires the user to respond within specific time may be inaccessible to users who are slow, for example users with physical or cognitive impairments. The application should notify the user of response time-out and allow the user to request additional time.

5.2. *Do not use flashing or blinking texts, objects or elements having a flash or blink frequency range of between 2 Hz and 55 Hz:* Flashing or blinking objects can result in photosensitive epileptic seizures in susceptive users.

5.5.4.1 Applicability of the IBM Software Accessibility Checklist to the Digital Doorway

Of the seventeen IBM accessibility guidelines, eleven are appropriate in formulating heuristics to evaluate direct accessibility support in the DD. The guidelines that are not directly applicable to the DD context are either aimed at applications providing indirect accessibility support through the use of assistive devices or they address functionalities that are not supported by the DD. Figure 5.11 provides a summary of the checklists that are relevant to the DD context.

1. All links, menus and control buttons should be accessible via the keyboard to allow flexibility in access to applications and the system's resources.
2. Provide a visible focus indicator that move among interface elements to draw users' attention to the current element.
3. Information should not be conveyed through images alone. A textual equivalent of the information should be provided.
4. A textual equivalent of sound alerts should also be provided so that the alert is not missed by people who are using the system in a noisy environment.
5. In situations where information is provided through a combination of audio and video, a synchronized textual equivalent should be provided with the primary information.
6. All interface elements should be given meaningful captions. Similar elements should be given the same captions across the application.
7. Users should be able to adjust the control of audio outputs.
8. The exclusive use of colour to provide information should be avoided.
9. There should be high contrast between interface elements and their background to enable access to people with low vision.
10. The font of instructions and captions should be large enough to enable access to people with low vision.
11. The use of automatic timed progression should be avoided. Where inputs have to be provided within a specific time, users should be notified of a response time-out and be given the opportunity to request for more time.

Figure 5.11: IBM software accessibility checklists applicable to Digital Doorway context

5.6 CONCLUSION

The main focus of this chapter was the design of accessible computer-based interactive systems. I started the chapter by giving some of the legal, ethical and economic reasons to make interactive systems accessible to people with disabilities.

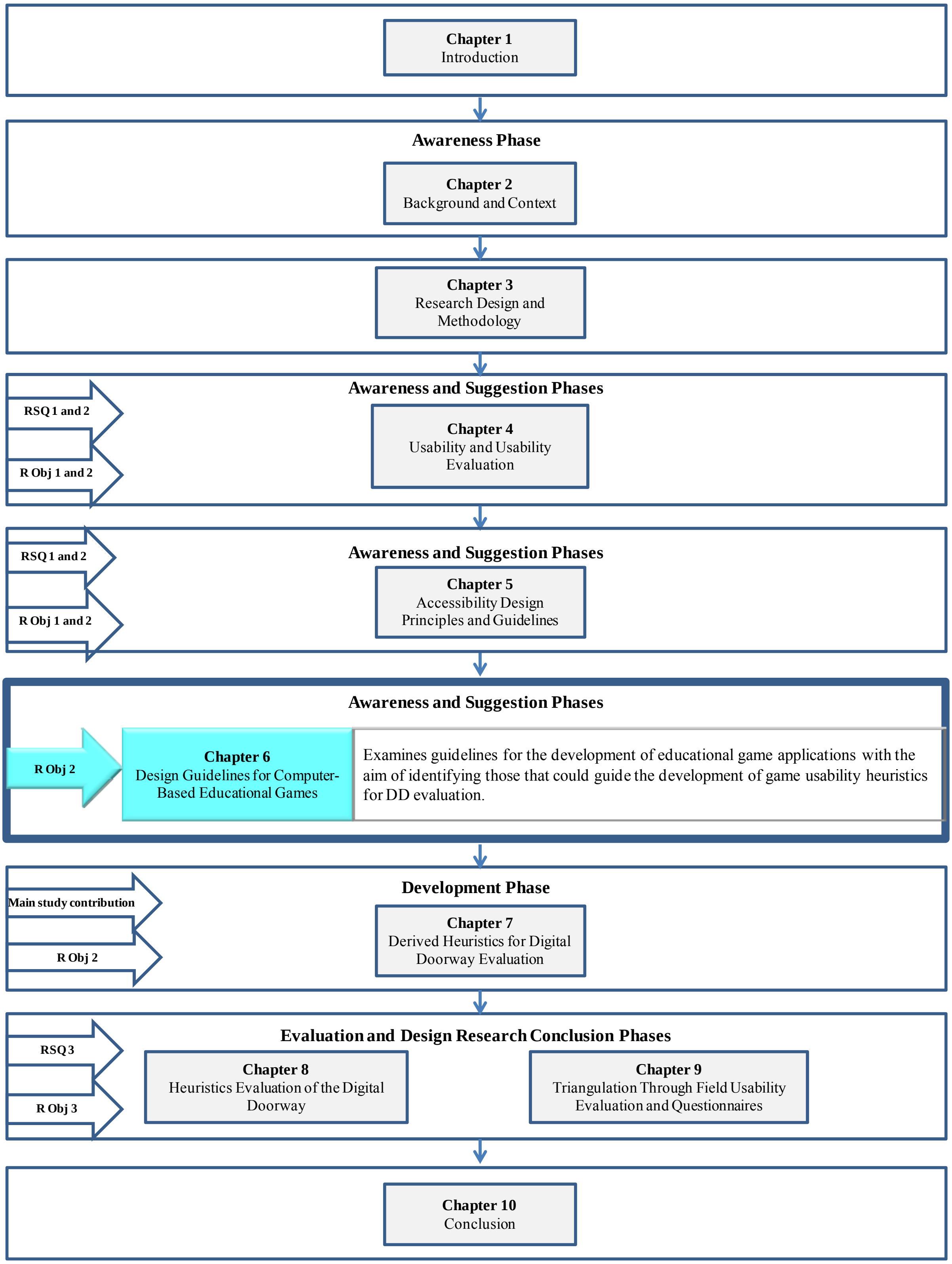
I also present an in-depth discussion of four categories of disabilities, the problems that are typically experienced by users with such disabilities and the various approaches for affording access to disabled users.

I looked at the techniques for evaluating the accessibility of interactive systems, with the aim of determining the appropriate methods for evaluating direct accessibility support provide in the DD. Design principles and guidelines for accessible interactive systems were also examined to assess their appropriateness in the development of heuristics for evaluating direct accessibility support in the DD.

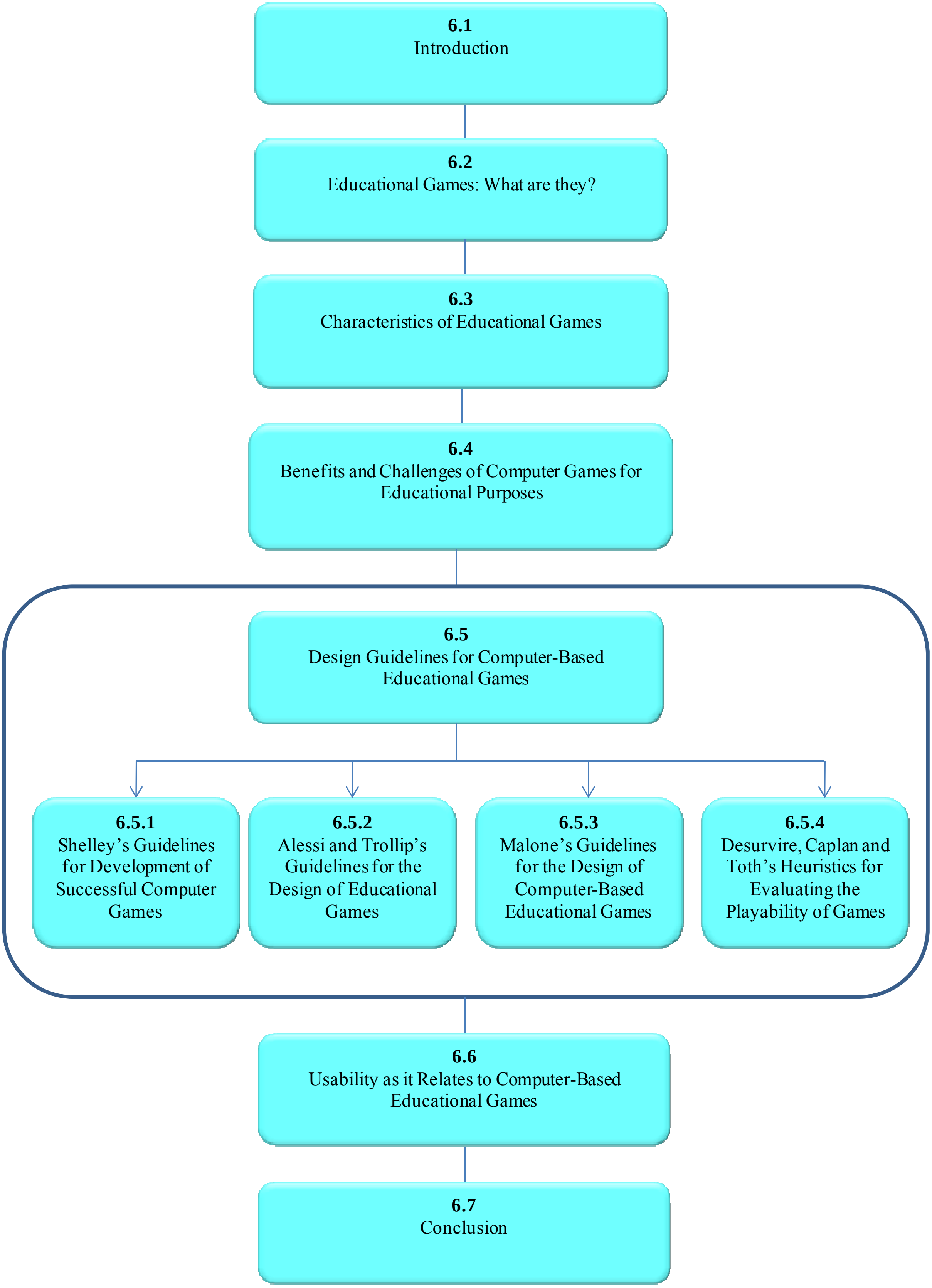
In Chapter 6, I will be discussing computer-based educational games and the principles guiding their design. Given that the three applications evaluated are educational games, it is important to examine the design guidelines for educational computer games. These design guidelines will inform the development of heuristics that will focus on the usability of the educational game applications selected for evaluation.

CHAPTER 6: DESIGN GUIDELINES FOR COMPUTER-BASED EDUCATIONAL GAMES

The stage of Chapter 6 in the dissertation



Map of Chapter 6



6.1 INTRODUCTION

This chapter maps onto the suggestion phase of the inner cycle of the design research process. It is concerned with the design guidelines for the development of educational games, a type of e-learning application. I begin the chapter with a brief description of what an educational game is (section 6.2). This is followed by an overview of the general features of educational games in section 6.3. In section 6.4 I discuss some of the benefits and challenges associated with the use of computer games for educational purposes. In section 6.5, I review the guidelines for the design of educational games proposed by Shelley [2001], Alessi and Trollip [2001], Malone [1980, 1981], and the heuristics for playability of games (HEP) developed by Desurvire et al. [2004]. In section 6.6 I look at what usability is in relation to educational game applications. Chapter conclusion is provided in section 6.7.

6.2 EDUCATIONAL GAMES: WHAT ARE THEY?

An educational game application is a type of e-learning application. E-learning is defined as "learning that is supported by information and communication technologies" [Cedefop, 2002:5-6]. E-learning is not restricted to the use of Internet for educational purposes, rather it include multiple formats and mixed methodologies, such as the use of software, CD-ROM, interactive multimedia applications, and online learning.

Educational games are programs designed for the acquisition of knowledge and skills (as opposed to mere entertainment) by incorporating elements of game, such as rules and competition. Learners solve problems by using the principles of the subject on which the application is based [Cruickshank and Telfer, 1980].

Educational games may be used for practice or to refine acquired skills, to identify weaknesses in knowledge or skills, to summarize or review a lesson, or to develop new relationships among concepts and principles [Gredler, 2004].

The use of games for educational purposes is not new. As early as the seventeenth century, war games were used by army and navy officials to improve their strategic planning for wars [Gredler, 2004]. Today, educational games range from adventure games, business games and those aimed at children for teaching arithmetic skills [Alessi and Trollip, 2001]. Games typically include specific rules and constraints guiding their use. The subject matter of this chapter is on the design principles guiding the development of educational games.

6.3 CHARACTERISTICS OF EDUCATIONAL GAMES

Educational games have certain features that distinguish them from other e-learning applications. In a literature analysis of 35 publications on educational video game design, Dondlinger [2007], noted the following characteristics of educational games:

- *Motivation:* This is an essential characteristic of educational games which propels learners to perform at their utmost abilities. Motivation can be intrinsic or extrinsic. Intrinsic motivation drive people to act freely on their own while extrinsic motivation forces one to act due to factors that are externally applied, for example reward or punishment [Denis and Jouvelot, 2005]. Intrinsic motivation is better than extrinsic motivation, where the reward (or punishment) is the main goal the learner strives for, and once this is removed the motivation disappears. Educational games can foster intrinsic motivation either by encouraging learners to master materials they would not ordinarily have chosen to learn or by enabling them to spend more time learning through the application than if the material was presented using other formats [Malone, 1981].

- *Narrative context:* The constructivist learning theory asserts that knowledge is influenced by a given social or cultural context. Educational games should provide a storyline that the player can identify with. The use of narratives can facilitate comprehension and recall [Waraich, 2004]. To be effective, there should be proper alignment between the learning content and the narrative context. This will allow learners to apply the desired knowledge and skills during their interaction with the application [Fisch, 2005].
- *Goals and rules:* The rules of the game define the role of players and the constraints imposed by the game. The rules must be mastered by players to attain their goals. To be effective, a game should have different levels of goals to sustain player motivation. Although goals and rules are typically integrated into the narrative context of a game, they are not subordinate to the context [Dondlinger, 2007; Malone, 1980, 1981].
- *Interactivity and multisensory cues:* Interactivity enables communication between the player(s) and the game environment. A highly interactive game engages the learner in meaningful activities. Learners can make choices, manipulate objects or test hypotheses, thereby enable them to experiment with different scenarios in a safe environment [Ritterfeld, Shen, Wang, Nocera and Wong, 2009]. The degree of user control affects the level of interactivity. A game with full user control has the tendency to become boring and unchallenging. In the same light, the player becomes more of a passive observer than an active participant in a game that imposes too much control. Hence, there should be a balance between the levels of user and program control [Swartout and Van Lent, 2003].

  Multisensory clues in the form of visual, auditory and haptic clues in a game environment can enhance the user experience. Multisensory displays or clues can focus the learner’s attention on important information and help prevent errors through feedback clues [Dondlinger, 2007].

6.4 BENEFITS AND CHALLENGES OF COMPUTER GAMES FOR EDUCATIONAL PURPOSES

Computer games provide learning environments that are both engaging and motivating [Fisch, 2005]. Learners can experiment with different scenarios which may be possible in real-life due to safety concerns [Corti, 2006]. The following paragraphs highlight some of the benefits and challenges to the use of educational games:

- *Learner motivation:* As stated in section 6.3, learning materials that are presented through the game approach can foster intrinsic learner motivation sine it can enthuse the use of materials which learners would not ordinarily have chosen and drive them to spend more time using the application [Alessi and Trollip, 2001].
- *Experimental learning:* Educational games that are combined with simulation enable learners to experiment with situations that may be too costly, risky or even impossible in the real world. For example, a business school student can discover the consequences of strategic business decisions through experimentation with an educational game [Alessi and Trollip, 2001; Corti, 2006; Pivec, 2007].
- *Foster cooperative learning and teamwork:* Educational games for multiple participants, where players support and help each other, can facilitate learning. The competitive nature of games make it possible to organize and match team members based on their distinctive capabilities [Alessi and Trollip, 2001; Bakar, Inal and Cagiltay, 2006].
- *Higher-order thinking skills:* The use of computer games for educational purposes can facilitate the development of higher-order thinking skills, such as critical thinking, problem solving, and decision making [Bakar et al., 2006]. These skills are supported through intrinsically motivating goals, fostering cooperation among learners and enabling experimentation and the application of prior knowledge into the game world.

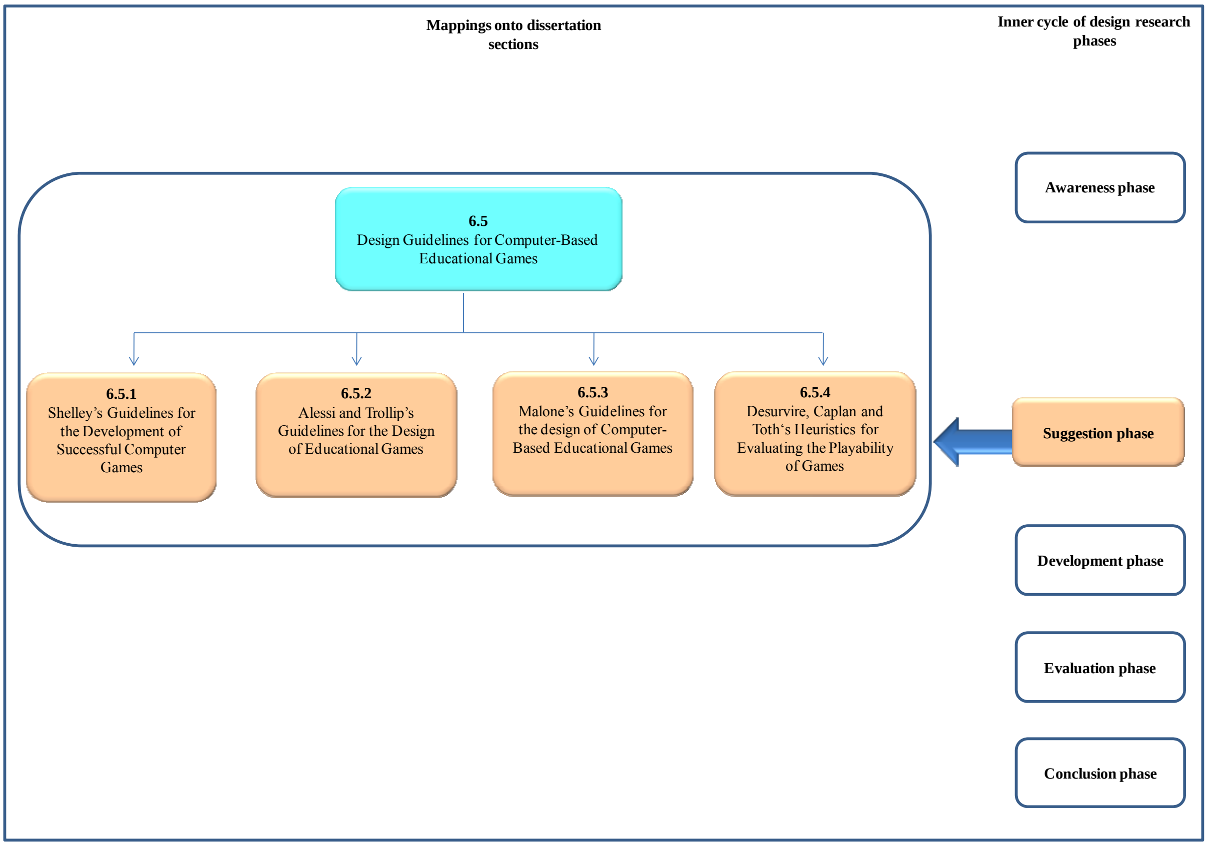
In spite of the benefits of educational games highlighted above, there are a number of challenges to their use, these include:

- *Creating games that are fun:* It is difficult to create an educational game that is fun to use by all potential users. This difficulty arises due to differences in human nature. What may be seen as a fun-filled game by a particular user may be perceived differently by others [Alessi and Trollip, 2001].
- *Reconciling the conflicts between educational and game objectives:* It is sometimes the case that the educational objectives conflict with the game objectives. One prominent area of conflict is in winning or losing a game. For example, the objective of winning a game may override the objective of learning by experimentation, where the learner avoids situations that may result in him/her losing the game [Alessi and Trollip, 2001].
- *Integrating educational content into the game environment:* Perhaps the main challenge to the use of computer games involves effective integration of the educational content into the game environment. This can be seen in the many drill-and-practice-like games which provide extrinsic motivation in the form of graphics and animations. The main problem with this approach is that learners may be drawn into the visual animations at the expense of the desired learning outcome [Fisch, 2005].

In the preceding sections, I have provided an overview of computer-based educational games, their main features, and the benefits and challenges associated with their use. Against this background, I can now better understand the special characteristics of the educational game applications on the DD. In the next section, I will discuss the guidelines that are specific for the design of educational games. This will assist me in developing evaluation heuristics that address the usability of the three educational game applications, *What-What Mzansi*, *OpenSpell* and *Themba's Journey*.

6.5 DESIGN GUIDELINES FOR COMPUTER-BASED EDUCATIONAL GAMES

Mapping of Section 6.5 the design research phases



Computer-based educational games are designed with the primary aim of facilitating the acquisition of new knowledge and skills, not just entertainment [Alessi and Trollip, 2001]. This section maps onto the suggestion phase of the inner cycle of the design research where I examine the guidelines for the design of educational games to determine their applicability to the DD context. First I discuss the design guidelines for computer games in general [Shelley, 2001], then I move to specific guidelines for educational games as proposed by Alessi and Trollip [2001] and Malone [1980, 1981]. The heuristics for evaluating the playability of games (HEP), proposed by Desurvire et al. [2004] are also examined in this section.

6.5.1 Shelley's Guidelines for Development of Successful Computer Games

The guidelines by Shelley [2001] were developed with commercial, entertainment games in mind. However, some of these guidelines are applicable to all games, whether educational or those for entertainment. Shelley's [2001] guidelines that can be applied to educational games are the following:

- *Reach for a broad audience:* Computer games should be designed for as wide user population as possible. Although different games appeal to different people, successful games attract both hardcore and casual gamers. Broad appeal can be achieved by ensuring that certain features of the game is engaging to the more frequent users while others appeal to the casual user.
- *Make the interface intuitive and easy to use:* The game interface is the first point of interaction between the user and the application. The interface should be intuitive, so that even casual gamers can begin meaningful interaction immediately. A confusing, difficult and frustrating interface may result in a player losing interest and never returning to the game.
- *Prototype early:* Game prototypes should be developed and tested early in the design process when it is much easier and less costly to make corrections. Early prototyping is essential for testing game play.
- *Design by playing:* As soon as a playable prototype is ready, designers should play it often to know what works and what doesn't. Designers should also include as many and diverse users as possible in the test play to create a game with wider appeal.
- *The player should have the fun, not the designer, programmer or the computer:* Computer games should be designed with the user in mind. Every aspect of game development should be directed at engaging the mind of the player. Every feature of the game, including graphics and sound effects should be directed at achieving the game objective.

6.5.1.1 Applicability of Shelley's guideline to the Digital Doorway

Two of the five computer game design guidelines by Shelley [2001] are not appropriate in the formulation of heuristics for evaluating the educational game applications in the DD. The inapplicable guidelines are:

- Early prototyping.
- Design by playing.

These guidelines focus on early phase of the development process. As stated in section 3.4.1.1, the DD is a fully functional computer system and the evaluation is summative.

Shelley's guidelines applicable to the DD are summarized in Figure 6.1.

1. *Reach for a broad audience:* Computer games should appeal to a diverse user audience. This guideline is particularly applicable to applications in the DD as the system is deployed to various communities across South Africa.
2. *Make the interface intuitive and easy to use:* The interface of a game application should be simple and intuitive so that users can begin meaningful interaction immediately.
3. *The player should have the fun, not the designer, programmer or the computer:* This guideline is similar in its interpretation to the general goal of user-centered design, where the focus is on the target users of the application.

Figure 6.1: Shelley's guidelines applicable to Digital Doorway context

6.5.2 Alessi and Trollip's Guidelines for the Design of Educational Games

Alessi and Trollip [2001] have developed educational-game specific guidelines. In their discussion on the essential guidelines for the design of educational games, Alessi and Trollip [2001] identify three typical sections in an educational game; the introduction; the body; and, the conclusion. The following sub-sections examine the three sections and the guidelines relevant to them.

6.5.2.1 Introduction of the Program

The introductory section of a game provides learners with the necessary information required for effective interaction. It typically begins with a title page, which provides general information about the application and authors' name. A title page should include control options for the learner to either continue or exit the program [Alessi and Trollip, 2001]. In addition to the title page, an introductory section should also include information regarding the following [Alessi and Trollip, 2001]:

- *Goal of the game:* A goal is the end to which every player strives for. Every game should have a goal that is either explicitly specified or easily inferred, for example scoring points. The goal of an educational game should reinforce the learning objectives.
- *Rules:* Rules specify what actions are allowed in the game and what constraints are imposed. Rules should be explicitly stated and easy to access throughout the program. Rules should also specify the role of each player, for example, in a multiplayer program; the requirement for special equipment, for example, a joystick); and, the applicable penalty for violation of game rules.
- *Directions:* These provide application-specific navigation instructions and how to set the level of difficulty. Directions should be given close to when they will be required and not at the beginning of the program so that learners do not forget them. Frequent users should be able to skip directions.

6.5.2.2 Body of the Game

This is the section where the central theme of the application, generally embedded in scenarios, is presented. The relevant guidelines for the design of game body that motivates learners and supports skills acquisition are [Alessi and Trollip, 2001]:

- *Challenge:* The level of challenge should be suited to the learner's capability and should not be too difficult or too easy. The level of difficulty can be selected by the learner or adjusted by the program as performance improves to maintain challenge throughout the game.
- *Curiosity:* This is closely related to challenge in that the process of satisfying one's curiosity can lead to overcoming a challenge. A high level of curiosity can motivate a learner to seek further information when conflicting ideas are presented. This is called cognitive curiosity [Malone, 1980, 1981].
- *Intended versus incidental learning outcome:* An educational game should facilitate both intended and incidental learning outcomes. Intended learning outcome arise when the skills and knowledge acquired in the game is closely related to the stated educational objectives. In incidental learning, the resulting skills are unrelated to the educational objectives.
- *Balance between skill and chance:* The game design should provide a good balance between the application of skill and sheer luck. Elements of luck will increase the level of surprise and challenge, but dependence on luck alone does not result in the attainment of learning objectives. At the same time, if the application of skill is just what is required to master the game, uncertainty is reduced and the game may lose its appeal.
- *Winning or losing the game:* A player can win a game by attaining specific goals or defeating an opponent (in multiplayer games). An improvement on previous performance may also be considered as winning. A game may be lost due to failure to achieve specified objectives. Losing a game should not demoralize or demotivate the learner.
- *Provision of information:* Information is required to enable the learner make informed choices and progress in the game. Some information should be provided at the beginning of the game, for example, rules and directions. Others are given as the game progresses, for example, feedback to learner actions. Depending on the nature of the game, the information may be accurate, partial, misleading or false. For instance in a detective game, witnesses may give false or misleading information, as is often the case in real life. When the application is likely to give false information, the learner must be aware of this possibility to avoid misconceptions regarding the computer's trustworthiness.
- *Locus of control:* This refers to whether the program or the learner is in control of the game. Ideally, a combination of program and learner control should be supported. Learner control of sequence, for example, forward and backward, as well as pace (how fast the application is) is important, as it can create a sense of being in control. Another important aspect of learner control is the ability to select the level of difficulty.

6.5.2.3 Concluding the Game

A game may be terminated temporarily by the learner (sometimes with the intention of returning to it later) or permanently when all required parts have been completed. Learners should not be forced to complete a game; hence there should always be an option to exit anywhere in the program. A request to temporarily terminate a game should be confirmed by the learner to prevent unintentional exit [Alessi and Trollip, 2001].

The following factors are relevant to the design of game conclusions [Alessi and Trollip, 2001]:

- *Recognizing the winner:* At the end of the game, the winning participant should be acknowledged. This may be done through verbal or visual messages. Some games store the points of top performers, for example, the top ten performers. In this case the score of the winner can be placed in the scoreboard if it ranks among the best. When scores are stored automatically, the player should be informed at the start of the game, especially if the player's real name is used to avoid embarrassing someone who might wish to remain anonymous. In addition to recognizing the winner, the program may also congratulate a player who has improved on previous performance to enhance learner motivation but this will require the storage of performance data from one session to another.
- *Providing information:* At the end of the game the learner should be given feedback regarding performance. Performance feedback should be positive and corrective. Negative and sarcastic comments should be avoided so as not to demoralize the learner.
- *Final message:* At the end of the game, there should be a clear message indicating that the program is ending. This will prevent a situation where the user is left wondering whether the game is indeed over.

6.5.2.4 Applicability of Alessi and Trollip's Educational Game Design Guidelines to the Digital Doorway

The three categories of guidelines by Alessi and Trollip [2001] can be applied in the formulation of heuristics for the DD.

- *Introduction to the game:* All educational game applications should have clear objectives; provide learners with the rules governing the game; and, navigation instructions. Hence, the three guidelines that relate to the design of the introductory part of educational game applications are relevant to the DD context.
- *Body of the game:* With regard to the guidelines on the design of the body of a game, some of these relate to content design, the evaluation of which is beyond the scope of this study. For this reason, the three guidelines "balancing of skills and luck"; "intended versus incidental outcome"; and, "curiosity" are excluded from the potential guidelines that could be used to formulate heuristics for evaluating the DD.

  Other guidelines that relate to appropriate level of challenge, provision of information and adequate control mechanisms are directly applicable to the DD context.
- *Game conclusion:* Only one of the three guidelines relating to the design of educational game applications' conclusion is applicable to the DD. The guidelines "recognizing the winner" is excluded since the game applications evaluated are single user applications and this guideline is

not pertinent to the usability of the applications. Likewise, the guideline on "final message" has been excluded because the three applications cannot be completed by users in a single session.

Figure 6.2 provides a summary of Alessi and Trollip's [2001] guidelines that can be used to derive evaluation heuristics for the DD.

1. *Goal of the game:* An educational game application should have clear goals and objectives.
2. *Rules:* The rules that govern the game, as well as the constraints that are imposed should be clearly specified.
3. *Directions:* Navigation instructions should be provided and easily accessible to learners.
4. *Challenge:* Provide an appropriate level of challenge based on the learner's capability.
5. *Provision of information:* Provide information on the performance of the learner in a constructive way.
6. *Locus of control:* Learners should be able to control how difficult the game should be. They should also have control over forward and backward progression of the game.

Figure 6.2: Alessi and Trollip's guidelines for educational games applicable to Digital Doorway context

6.5.3 Malone's Guidelines for the Design of Computer-Based Educational Games

In his guidelines for the design of computer-based educational games, Malone [1980, 1981] provides a three-category taxonomy of intrinsic motivators that make learning activities fun. These are challenge, fantasy and curiosity. This sub-section reviews the three categories, together with the guidelines that fall under each category.

6.5.3.1 Challenge

A computer game can only be challenging if the attainment of its goal is uncertain [Malone, 1980]. A game becomes boring if the player is sure of his/her ability to win. There are four different ways to provide an uncertain outcome in a game [Malone, 1980, 1981]:

1. *Variable level of difficulty:* To provide adequate challenge, a game can have different levels of difficulty determined by the program, set by the player, or determined by the skills of the opponent. Varying the level of difficulty ensures that the learner is not demotivated right from the beginning because s/he can choose to start at an appropriate level.
2. *Multiple levels of goals:* The outcome of a game can be uncertain if there are at least two levels of goals (i) attaining the basic outcome of the game, for example, solving an arithmetic problem and (ii) achieving the basic goal efficiently, for instance, by solving the problem quickly.
3. *Hidden information:* Uncertainty can be implemented by hiding information required to attain the goal or selectively revealing it. This provides challenge and increases the player's curiosity.

4. *Randomness:* The outcome of a game can be made uncertain by incorporating elements of randomness or chance into it.

Challenge is an important intrinsic motivator as accomplishment can improve the leaner's self-esteem. Likewise, failure has the potential to lower self-esteem, even to the extent of discouraging the leaner from playing the game again. Providing performance feedback that does not demean the leaner will reduce the possibility of damaging a leaner's self-esteem [Malone, 1980, 1981].

6.5.3.2 Fantasy

Fantasy games are interesting, evoking imaginary images or a close representation of reality. Fantasies can be extrinsic or intrinsic. Extrinsic fantasies are applicable to different subject areas, and the fantasy depends on the correct use of a skill. For example, a player may try to avoid some fantasy catastrophe by guessing the correct word in a guessing game [Malone, 1980, 1981]. In intrinsic fantasies, both the skill that is being learnt and the fantasy world are dependent on each other. The scenarios in the game depend on the correct use of the skill and the way skill usage varies from the correct or expected way of use. According to Malone [1980, 1981] intrinsic fantasies are more interesting and educational than extrinsic fantasies because they show how the skills acquired can be applied in real world. They enable learners exploit analogies, where existing knowledge is applied to the new fantasy world. Furthermore, provoking vivid images of materials learnt can facilitate remembrance [Malone, 1980, 1981].

6.5.3.3 Curiosity

Curiosity is an intrinsic motivator which compels people to learn beyond their current knowledge. Computer-based educational games should provide "optimal level of informational complexity, being neither too easy nor too complex with respect to the learner's existing knowledge" [Malone, 1981:362]. Learners should have sufficient knowledge to foresee that some events will happen though some of the events may never come to pass [Malone, 1981].

Malone [1980, 1981] makes distinctions between sensory and cognitive curiosity. Sensory curiosity is aroused through the use of attention-grabbing presentation and multi-media effects, such as colour, graphics and sound. This helps to sustain the learner's attention in the presentation. Multi-media effects can be used for decoration, enhance fantasy, for reward, and as system representation to convey information in an effective way.

Cognitive curiosity is aroused by the need to modify or correct incomplete or inconsistent information. Malone [1980, 1981] gives an example where learners seek to fulfil their cognitive curiosity following a seemingly contradictory Biology lesson, where they were told that plants require sunlight for the process of photosynthesis but that some plants, like fungi, can live in darkness.

6.5.3.4 Applicability of Malone's Educational Game Design Guidelines to the Digital Doorway

Only one of the three categories of the intrinsic motivators proposed by Malone [1980, 1981] is directly applicable to the DD context. The inapplicable guideline categories are:

- *Fantasy:* The three educational game applications evaluated are not fantasy games; hence this guideline has been excluded from potential guidelines that could be used to formulate heuristics for evaluating the DD.
- *Cognitive curiosity:* As stated in section 6.5.2.4, the guideline on curiosity relates to content design and their evaluation is beyond the scope of this study. The guideline is thus excluded.

The only applicable guideline category is:

- *Challenge:* Malone's [1980, 1981] guideline regarding the provision of appropriate level of challenge is similar in meaning and interpretation to that by Alessi and Trollip [2001]. Likewise, the provision of performance feedback in such a way that learners are not demotivated is important and relevant for the evaluation of educational game applications in the DD.

A summary of Malone's [1980, 1981] guidelines that could guide the development of heuristics to evaluate the usability of educational game applications in the DD is provided in Figure 6.3.

1. Educational game applications should neither be too simple nor too complex.
2. Provide support for varying levels of difficulty, adjustable by the application or the learner.
3. Performance feedback should be constructive.
4. Performance feedback should not demean learners.

Figure 6.3 Malone's educational game design guidelines applicable to Digital Doorway context

6.5.4 Desurvire, Caplan and Toth's Heuristics for Evaluating the Playability of Games

The heuristics for playability of games (HEP), proposed by Desurvire et al. [2004], were based on literature study of heuristics for playtesting of games and productivity applications. To assess the effectiveness of the heuristics, playability evaluators and game designers were asked to evaluate a game application at the design stage. The results from the expert evaluation were then validated and compared with the problems identified through user testing.

HEP heuristics consists of four categories: *game play*; *game story*; *game mechanics*; and, *game usability* heuristics. Game play relates to the problems and challenges that a player must overcome to win the game. Game story includes all the plot and character development and game mechanics deals with the programming that provides the structure through which the various units interact with the game environment. Game usability relates to the game interface and the controls through which the player interacts with the application [Desurvire et al., 2004]. HEP is particularly useful in evaluating game applications during design stage. This does not imply that it cannot be used to evaluate a fully functional game application. Section 6.5.4.1 examines the game usability heuristics category and their applicability to the DD context.

6.5.4.1 The Game Usability Heuristics Category of HEP

The game usability category of HEP consists of twelve heuristics [Desurvire et al., 2004]:

1. Provide immediate feedback for user actions.
2. The player can easily turn the game off and on, and be able to save games in different states.
3. The player experiences the user interface as consistent (in control, colour, typography, and dialogue design) but the game play is varied.
4. The player should experience the menu as a part of the game.
5. Upon initially turning the game on, the player has enough information to get started to play.
6. Players should be given context sensitive help while playing so that they do not get stuck or have to rely on a manual.
7. Sounds from the game provide meaningful feedback or stir a particular emotion.
8. Players do not need to use a manual to play game.
9. The interface should be as non-intrusive to the player as possible.
10. Make the menu layers well-organized and minimalist to the extent the menu options are intuitive.
11. Get the player involved quickly and easily with tutorials and/or progressive or adjustable difficulty levels.
12. Art should be recognizable to player, and speak to its function.

6.5.4.2 Applicability of Desurvire et al.'s Game Usability Heuristics to Digital Doorway

The game usability heuristics category of (HEP) focuses on the usability of game interfaces and could be useful in evaluating the educational game applications *What-What Mzansi*, *OpenSpell* and *Themba's Journey*.

Four of the heuristics were excluded from those that could be used to formulate heuristics for evaluating the DD. The heuristics "the player should experience the menu as a part of the game" and "art should be recognizable to player and speak to its function" are excluded because their meaning and purpose are not clear. The heuristic "players do not need to use a manual to play game" is excluded since the DD is used without any user manual, while the heuristic "the interface should be as non-intrusive to the player as possible" is excluded because the game applications evaluated are simple educational games that do not require user immersion, which are sometimes used in three-dimensional interfaces.

The specific heuristics that are directly applicable to DD context are summarized in Figure 6.4.

1. Provide immediate feedback for user actions. Feedback to user action is essential for any interactive system, including game applications. Without feedback, users cannot determine the effect of actions that have been initiated and it becomes difficult to proceed with the interaction between the user and the application.
2. There should be appropriate control mechanisms to support easy navigation by users.
3. The player should experience the user interface as being consistent, though the game play can be

varied. Consistencies in the way similar actions are initiated are especially important in the context of the DD. Because the DD is aimed at users with little or no computer literacy, interface consistency will allow the transfer knowledge from one application to similar ones.

4. Upon initially turning the game on, the player should have enough information to get started. Information regarding game objectives, rules and any constrain imposed by the game application will enable the user begin meaningful interaction immediately.
5. Meaningful feedback should be provided.
6. Players should be given context sensitive help while playing so that they do not get stuck. Help information that is based on the current activity which a player is involved in will allow the player to associate the help information with the activity. Help information that is provided in a decontextualized way may go unnoticed by the player.
7. Menu options should be simple, with meaningful and intuitive labels.
8. Provide an adjustable difficulty level to ensure constant challenge to the player and reduce boredom.

Figure 6.4 HEP game usability heuristcis applicable to Digital Doorway context

In this section, I have examined the general guidelines for the design of computer games and the guidelines that are specific to educational game applications. Since my study involved the usability evaluation of three educational game applications installed on the DD, in the next section, I will briefly discuss how usability in this context differ from other applications like transaction processing applications.

6.6 USABILITY AS IT RELATES TO COMPUTER-BASED EDUCATIONAL GAMES

An educational game application is a form of e-learning application used to foster learner motivation and provide a safe environment for experimentation [Alessi and Trollip, 2001; Corti, 2006; Pivec, 2007]. In contrast to a transaction processing system that aims for rapid completion of tasks, the primary goal of educational applications is to support learning [Ardito, Costabile, De Marsico, Lanzilotti, Levialdi, Roselli and Rossano, 2006].

The usability of educational applications is essential since users typically have to deal with double learning processes, learning to use the system and learning the content presented [Parlangeli, Marchigiani and Bagnara, 1999]. The users of a computer-based educational game should not spend substantial amounts of time learning how to use the application. The interface should be intuitive, so that even novice users can begin meaningful interaction immediately [Ardito et al., 2006; Desurvire et al., 2004].

Educational applications should be motivating, provide appropriate level of learner challenge, and constructive feedback.

The interface should neither confuse nor distract the learner [Alessi and Trollip, 2001; Ardito et al., 2006; Malone, 1980, 1981]. Interface usability errors that distracts from learning should be prevented. However, cognitive errors, which form part of the learning process, should be permitted with adequate support for the recognition, diagnosis, and recovery from these types of errors [Squires and Preece, 1999].

6.7 CONCLUSION

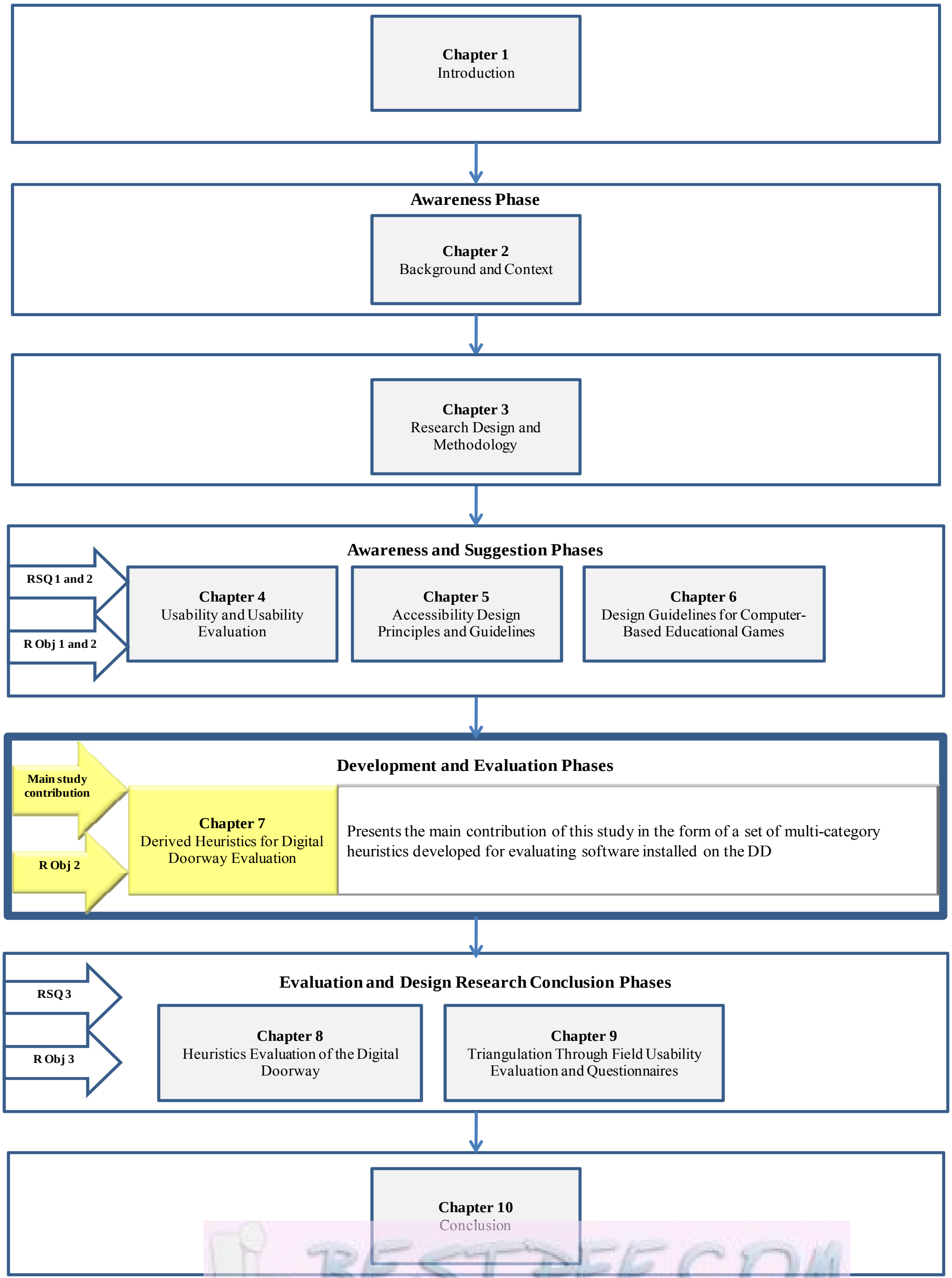
The primary purpose of this chapter was to examine the principles and guidelines for the design of computer-based educational games to determine their appropriateness for the development of evaluation heuristics for DD evaluation.

To set the context, I provide the description of what is regarded as an educational game. This was followed by an overview of general characteristics of educational games, and the benefits and challenges to their use. Guidelines for the design of educational game applications were also examined. A brief discussion of what is involved in the usability of educational game applications was also provided.

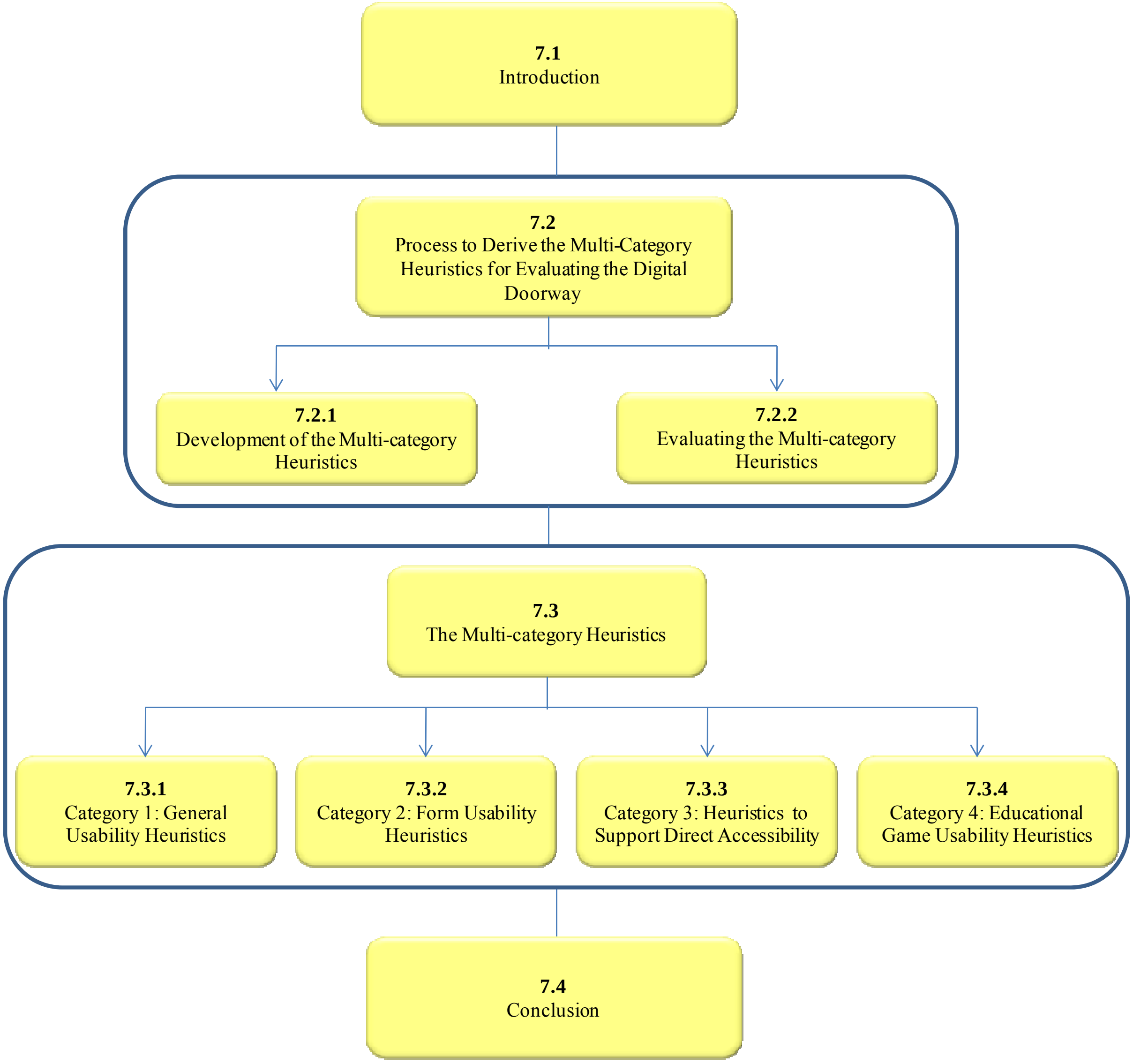
In Chapter 7, I will present the main contribution of this study, namely, the evaluation heuristics that emerged from the literature investigation of principles and guidelines for the design of usable and accessible interactive systems, including educational computer games.

CHAPTER 7: DERIVED HEURISTICS FOR DIGITAL DOORWAY EVALUATION

The stage of Chapter 7 in the dissertation



Map of Chapter 7



7.1 INTRODUCTION

This chapter maps onto the development and evaluation phases of the inner cycle of the design research process. It is directly linked to my research objective 2: To develop an instrument that can be used to evaluate the usability and direct-accessibility support provided in the DD. It provides the main contribution of the study, the set of multi-category heuristics specifically developed for evaluating the usability and direct accessibility support provided by the DD. The development of the multi-category heuristics is necessitated after the heuristic evaluation method was identified as an appropriate method to evaluate a selection of interfaces and applications on the DD, in answer to my research sub-question 2: Which of these methods can be applied in the evaluation of the DD?

The heuristics are based on the design principles, guidelines and heuristics examined in Chapters 4, 5, and 6. In section 7.2, I discuss how the set of multi-category heuristics was developed and evaluated for its completeness. The heuristics are presented in section 7.3 and chapter conclusion is provided in section 7.4.

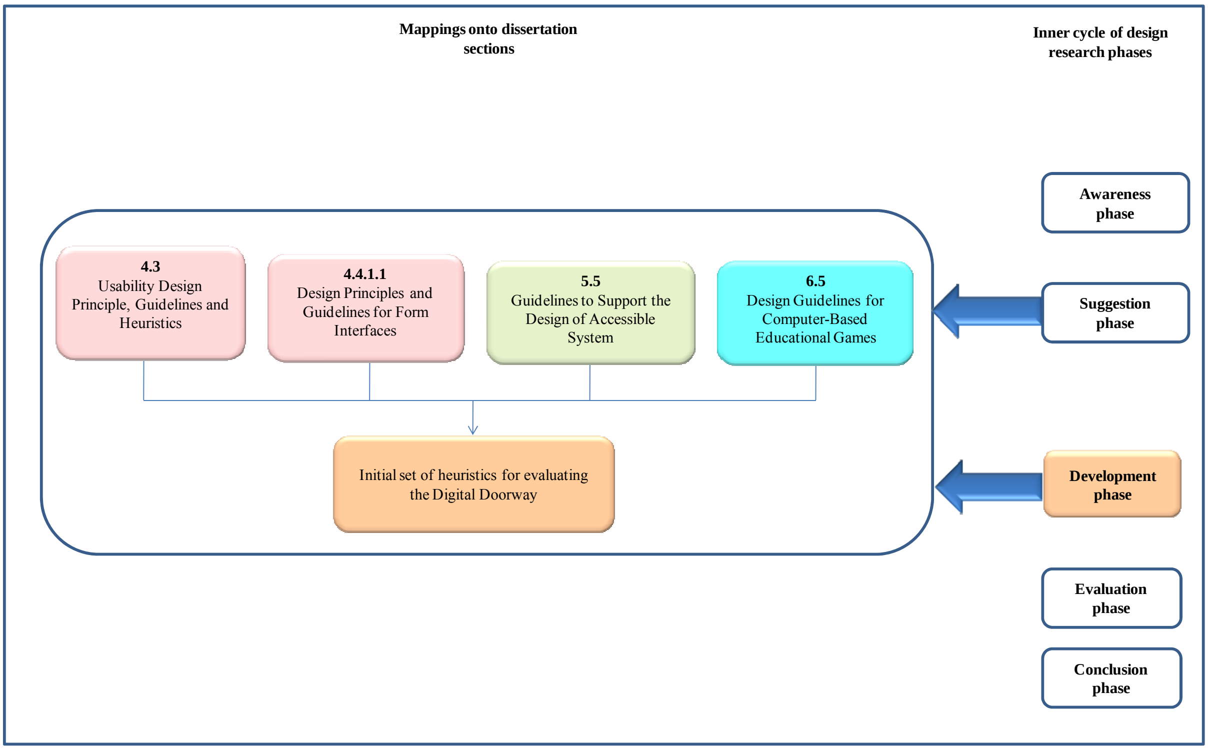
7.2 PROCESS TO DERIVE THE MULTI-CATEGORY HEURISTICS FOR EVALUATING THE DIGITAL DOORWAY

Heuristic evaluation method is suitable for formative and summative evaluation of interactive systems (see section 4.2.3.1). To be of any value, the set of heuristics must be appropriate for the specific application being assessed [Jeffries et al., 1991; Preece et al., 2007].

A detailed discussion of the process followed to develop the application-specific heuristics for evaluating the DD is given in section 7.2.1 while section 7.2.2 provides the description of how the heuristics were evaluated.

7.2.1 Development of the Multi-category Heuristics

Mapping of section 7.2.1 to the design research phases



This section maps onto the development phase of the inner cycle of the design research process. It follows on the suggestion phase of the inner cycle by presenting the process followed to derive the heuristics for evaluating the interfaces and applications on the DD.

The DD is aimed at users with little or no computer literacy, and potential users may not have seen a computer previously. General usability principles, such as those relating to intuitive interfaces, the provision of adequate feedback and consistency in the way similar actions are initiated will ensure that users can begin effective interaction with the DD with as little difficulty as possible.

Because the evaluation is focused on direct-accessibility support, in addition to the general usability of a selection of interfaces and applications on the DD, it was important that the set of heuristics addressed the two concerns.

Usability principles and guidelines, such as those put forward by Dix et al. [2004] and Nielsen [1994b], are primarily aimed at general usability of interfaces. Such principles may not be appropriate or sufficient for evaluating direct accessibility support. Furthermore, because the DD applications that are evaluated are also educational games, the usability of which may not be covered by general usability guidelines, a combination of principles and guidelines that address general usability, accessibility and educational game usability is required to derive an appropriate set of heuristics for the evaluation.

In Chapter 4 (sections 4.3 and 4.4.1.1), Chapter 5 (section 5.5) and Chapter 6 (section 6.5), which map onto the suggestion phase of the inner cycle of the design research, I examined a number of design principles, guidelines and heuristics for their applicability to the DD context. At the end of the specific sections, I highlighted, in shaded blocks, the principles, guidelines and heuristics that were found to be appropriate to derive evaluation heuristics for the DD. The rationale for the inclusion and exclusion of the principles and guidelines were provided in those sections.

To determine their applicability or non-applicability to DD context, each of the principles and guidelines were examined by considering the interfaces and applications to be evaluated, the types of users the DD is aimed at, and the typical environment of DD usage.

As an example, eleven of the usability principles of Dix et al. [2004], discussed in sections 4.3.1.1 to 4.3.1.3, were found to be applicable to the DD context. In section 4.3.1.4, the applicability of these principles was discussed. All five principles under learnability were found to be applicable to the DD context, since they address the ease with which users are able to learn and use new systems.

Examples of principles excluded include the principle of *multi-threading*, because the DD only supports the execution of one task at a time, and *task migratability*, since the applications that will be evaluated do not require the transfer of control for task execution between the applications and the user.

A similar process was followed for the remaining general usability principles and guidelines, the form design guidelines, accessibility principles and guidelines, and educational game design guidelines.

To aid readability and facilitate the analysis of identified problems, the heuristics are grouped into four categories: general usability heuristics; form usability heuristics; direct accessibility heuristics; and, educational game usability heuristics:

1. *General usability heuristics:* The derived heuristics, focussing on general interface usability, are based on the usability design principles proposed by Dix et al. [2004], guidelines for the design of children's technology of Gelderblom [2008], the general interface heuristics of Nielsen [1994b], design principles of Norman [2001], usability and user experience goals proposed by Preece et al.

[2007], and design golden rules of Shneiderman [1998]. Other principles and guidelines that provide input for this heuristic category are the game interface usability heuristics of Desurvire et al. [2004] and Shelley's [2001] guidelines for design of successful game applications.

2. *Form heuristics:* One of the interfaces evaluated is an electronic form. Forms have specific guidelines to ensure they are usable by the target user group. The heuristics focusing on form interface usability and accessibility are based on the form interface design principles and guidelines proposed by Mayhew [1992], usability principles of Dix et al. [2004], the guidelines for the design of children's technology of Gelderblom [2008], the accessibility checklist proposed by IBM [2009] and the United State's electronic and information technology accessibility standards [2000].
3. *Direct accessibility heuristics:* To provide adequate coverage for direct accessibility issues, design principles and guidelines with specific focus on interface accessibility were examined. Accessibility principles and guidelines reviewed were the universal usability principles of Story et al. [1998], the WCAG 1.0 [1999], the United State's electronic and information technology accessibility standards [2000], and the software accessibility checklist proposed by IBM [2009].
4. *Game heuristics:* To address the special requirements of educational game usability, the design guidelines for successful computer game applications put forward by Shelley [2001], the educational game design guidelines of Alessi and Trollip [2001] and Malone [1980, 1981], and the game usability heuristics proposed by Desurvire et al. [2004] were examined.

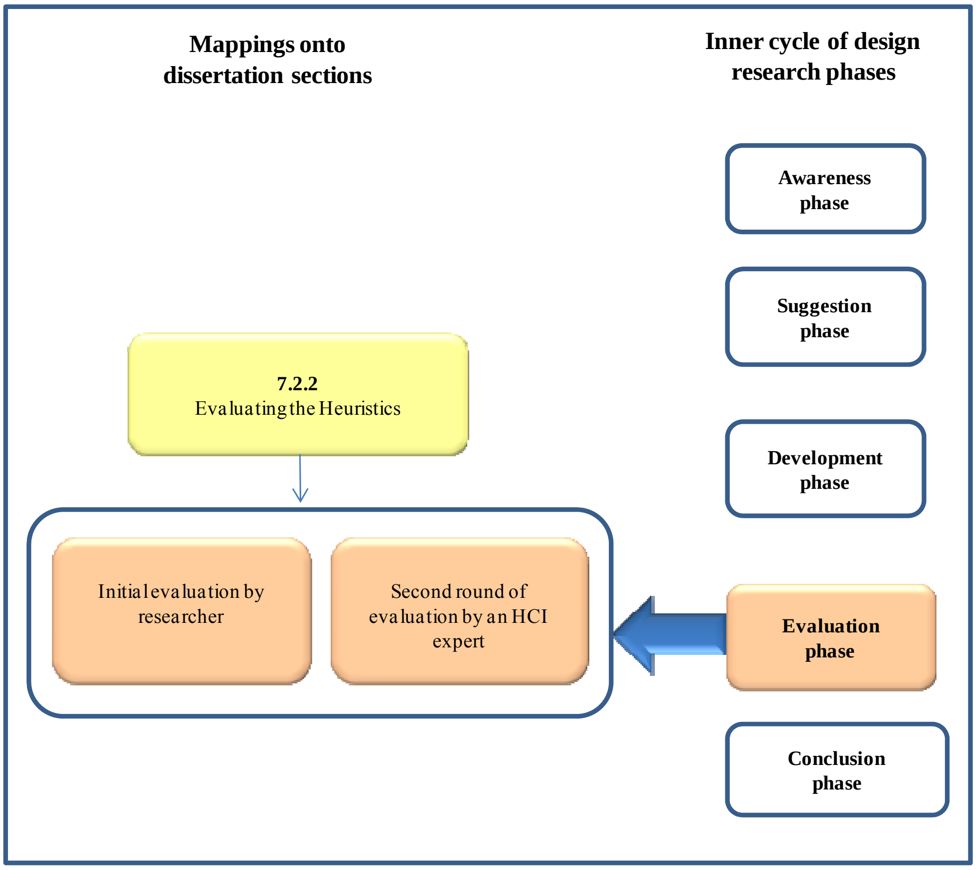
Within each category, the applicable heuristics were further grouped into high level heuristics following the analysis of the design principles, guidelines and heuristics reviewed in sections 4.3, 4.4.1.1, 5.5, and section 6.5:

- The high level groupings in the general usability heuristic category were guided by the main themes of the design principles proposed by Dix et al. [2004] and Nielsen's heuristics.
- The groupings in the form usability heuristic category were guided by the main issues covered in the form interface design guidelines of Mayhew [1992].
- Those in the direct accessibility heuristic category were based on the primary focus of the universal design principles of Story et al. [1998] and IBM [2009] software accessibility checklist.
- The main themes of the educational game design guidelines proposed by Alessi and Trollip [2001] and the general guidelines for game applications put forward by Shelly [2001] guided the groupings provided in the educational game usability heuristic category.

The high level heuristics in each category provide specific sub-heuristics to facilitate the identification of usability and/or direct accessibility problems. The outcome of this process was an initial set of multi-category evaluation heuristics.

7.2.2 Evaluating the Multi-category Heuristics

Mapping of section 7.2.2 to the design research phases



This section maps onto the evaluation phase of the inner cycle of the design research. The process followed to evaluate the multi-category heuristics is discussed in the following sub-sections.

7.2.2.1 Initial Evaluation of the Heuristics

At the evaluation phase of the inner cycle, the initial set of heuristics were tested on the selection of interfaces and applications on the DD. Using the heuristics, I conducted heuristic evaluation on the DD login screen, the registration form, the main desktop, and the three educational game applications *What-What Mzansi*, *OpenSpell* and *Themba's Journey* to assess whether the heuristics provide adequate coverage of potential usability and direct accessibility problems in these interfaces and applications.

After the initial evaluation, the identified problems were matched against the initial set of multi-category heuristics. The heuristics were then modified to provide for the identified problems that could not be matched to any of the heuristics.

7.2.2.2 Second Round of Evaluation

Following the initial round of evaluation and modifications to the heuristics, an HCI expert, who had previous experience in usability and accessibility issues, conducted another round of heuristic evaluation on the selected interfaces and applications using the modified heuristics. This expert was not included in the team of five experts who conducted the later formal heuristic evaluation on the DD (see section 7.2.2.3). The aim of the evaluation by the HCI expert was also to assess the completeness and terminology usage in the heuristics. Further modifications were made to the heuristic sets based on the suggestions by the expert. The outcome of this process was the final set of multi-category heuristics; this set will be presented in section 7.3.

7.2.2.3 Formal Heuristic Evaluation using the Multi-category Heuristics

As part of the evaluation activities in the outer cycle of the design research process, a formal heuristic evaluation was conducted on the selected interfaces and applications by a team of five evaluators using the final set of multi-category heuristics.

Following the five heuristic evaluation sessions, a large number of usability and direct accessibility-related problems were identified by the evaluators. Problems that were of similar nature were consolidated to provide an aggregated list of 71 usability and direct accessibility problems. Section 8.3.1 discusses the heuristic evaluation process in detail.

On completion of this process, two of the problems identified by the experts could not be matched to any of the proposed heuristics. One of these problems related to the absence of a mechanism to retrieve a forgotten password on the login screen. The other involved the inability of users to temporarily exit at any section in *Themba's Journey*, which requires a substantial amount of time to complete, and return to the same section at a later stage. The following additional heuristics were generated to cover these problems:

1. Follow and adhere to platform and industry standards and conventions.
2. Whenever appropriate, give users the option of returning to where they left off when the program is temporarily exited.

7.3 THE MULTI-CATEGORY HEURISTICS

The outcome of the processes discussed in sections 7.2.1 to 7.2.2 was the set of multi-category heuristics suitable for evaluating the usability and direct accessibility support provided by the DD. As stated in section 7.2.1, the heuristics were grouped into four categories, namely, general usability, form usability, direct accessibility, and educational game usability heuristics. This is to aid their readability and facilitate the analysis of identified problems.

The general usability and direct accessibility heuristics are applicable to all interfaces and applications. The form usability heuristics are targeted at the registration form while educational game heuristics are related to *What-What Mzansi*, *OpenSpell* and *Themba's Journey*. Sections 7.3.1 to 7.3.4 describe the four categories of evaluation heuristics while Tables 7.1 to 7.4 provide the heuristics that belong to each category.

Within each heuristic category there are a number of high level heuristics (shown in boldface in Tables 7.1 to 7.4) and specific sub-heuristics to guide evaluation. The heuristics are numbered in order for easy referencing. For example, within the general usability heuristics category, the sub-heuristic "provide information that will enable users understand how to interact with the Digital

Doorway using clear and simple terminology" is numbered 1.1 in column 1 (H_No). The descriptions of the heuristics are given in column 2 (Heuristics), while the specific sections in this dissertation where the heuristics have been discussed are shown in column 3 (Applicable sections). Column 4 (Reference) provides references to the sources of the principles and guidelines on which the heuristics are based.

7.3.1 Category 1: General Usability Heuristics

The principles, guidelines and heuristics on which Category 1 heuristics are based are provided in section 7.2.1. There are ten high level evaluation heuristics within this category (see also section 7.2.1 for the basis of the high level groupings). Each of the heuristics has specific sub-heuristics for evaluation, with the total number of specific sub-heuristics in the category being 29. Table 7.1 describes the general usability heuristics for the DD.

Table 7.1: General usability heuristics for Digital Doorway evaluation

H_No	Heuristics	Applicable sections	Reference
1	**Support user efforts to learn and use the Digital Doorway**		
1.1	Provide information that will enable users understand how to interact with the Digital Doorway using clear and simple terminology.	4.3.1.1 4.3.2 4.3.4.1	Dix et al. [2004] Gelderblom [2008] Preece et al. [2007]
1.2	Provide clear indication of what the next required action is.	4.3.1.1 4.3.2 4.3.4.1	Dix et al. [2004] Gelderblom [2008] Preece et al. [2007]
2	**Support users in the application of other computer or real world experiences while interacting with the Digital Doorway**		
2.1	Avoid the use of technical terms.	4.3.3	Nielsen [1994b]
2.2	Icons, symbols and menu items should be labelled with intuitive and meaningful names, taking into account user context and experience.	4.3.1.1 4.3.2 6.5.1 6.5.4	Dix et al. [2004] Gelderblom [2008] Shelley [2001] Desurvire et al. [2004]
2.3	Ensure that information sequence and layout appear in natural and logical order.	4.3.3	Nielsen [1994b]
2.4	Follow and adhere to platform and industry standards and conventions.	4.3.3	Nielsen [1994b]
2.5	The mappings between controls and their effect should be intuitive and easily understood.	4.3.2 4.3.5	Gelderblom [2008] Norman [2001]
3	**Ensure that users' short-term memory is not overloaded**		
3.1	Users should not use considerable cognitive resources trying to interpret the meaning of icons, menus and symbols, and to navigate the interface.	4.3.2 4.3.3 4.3.6	Gelderblom [2008] Nielsen [1994b] Shneiderman [1998]
3.2	Objects, options and permissible actions should be visible so that users do not have to remember instructions.	4.3.2 4.3.3 4.3.5	Gelderblom [2008] Nielsen [1994b] Norman [2001]

3.3	Users should not be required to remember information from one screen to another.	4.3.2 4.3.3	Gelderblom [2008] Nielsen [1994b]
3.4	Menu options should be logically grouped to aid the recognition of available functionalities.	4.3.1.1 4.3.2 4.3.3 4.3.4 4.3.5 4.3.6	Dix et al. [2004] Gelderblom [2008] Nielsen [1994b] Preece et al. [2007] Norman [2001] Shneiderman [1998]
4	**Provide observable and informative feedback regarding change in system state**		
4.1	Feedback should be provided in clear and unambiguous terms.	4.3.1.1 4.3.2 4.3.3 4.3.5 4.3.6 6.5.4	Dix et al. [2004] Gelderblom [2008] Nielsen [1994b] Norman [2001] Shneiderman [1998] Desurvire et al. [2004]
4.2	Any change in the state of the system following user action should be perceivable to enable the user associate the change to the action that caused it.	4.3.1.1 4.3.2 4.3.3 4.3.5 4.3.6 6.5.4	Dix et al. [2004] Gelderblom [2008] Nielsen [1994b] Norman [2001] Shneiderman [1998] Desurvire et al. [2004]
4.3	Response to user action by the system should be instantaneous. Where this is not possible, the system should indicate that the task is in progress to avoid repeated clicking by the user.	4.3.1.3 4.3.2 4.3.4.1 4.3.6	Dix et al. [2004] Gelderblom [2008] Preece et al. [2007] Shneiderman [1998]
5	**The application should be tolerant of users' mistakes**		
5.1	Provide support for system exploration by the user by allowing easy reversal of actions.	4.3.1.3 4.3.3.1 4.3.4.1 4.3.6 5.5.1.1	Dix et al. [2004] Nielsen [1994b] Preece et al. [2007] Shneiderman [1998] Story et al. [1998]
5.2	Prevent user error by using appropriate constraints at strategic points.	4.3.2 4.3.4.1 4.3.5 4.3.6 5.5.1.1	Gelderblom [2008] Preece et al. [2007] Norman [2001] Shneiderman [1998] Story et al. [1998]
6	**Help users recognize, diagnose and recover from errors**		
6.1	Error messages should be context-specific in relation to the action performed.	4.3.2 4.3.3 6.5.4	Gelderblom [2008] Nielsen [1994b] Desurvire et al. [2004]
6.2	Error messages should be given in language comprehensible to users, not using technical terms.	4.3.2 4.3.3	Gelderblom [2008] Nielsen [1994b]
6.3	Error messages should precisely describe what the problem is and offer ways of solving them.	4.3.2 4.3.3	Gelderblom [2008] Nielsen [1994b]
7	**Keep interface elements simple through minimalist design**		

7.1	Ensure that the interface is not cluttered with irrelevant information, control buttons and icons.	4.3.3 6.5.4	Nielsen [1994b] Desurvire et al. [2004]
7.2	Provide information and control options close to when the user is required to make use of them.	4.3.2	Gelderblom [2008]
8	**Ensure internal consistency within and across Digital Doorway applications**		
8.1	Be consistent in the naming conventions used for icons, symbols and objects.	4.3.1.1 4.3.3 4.3.6 6.5.4	Dix et al. [2004] Nielsen [1994b] Shneiderman [1998] Desurvire et al. [2004]
8.2	Make sure that the same terms, actions or symbols mean the same thing across applications.	4.3.1.1 4.3.3 4.3.6 6.5.4	Dix et al. [2004] Nielsen [1994b] Shneiderman [1998] Desurvire et al. [2004]
8.3	Create the same 'look and feel' effect across applications so users can extend knowledge to similar situations.	4.3.1.1 4.3.3	Dix et al. [2004] Nielsen [1994b]
9	**Match between component labels and their content**		
9.1	Ensure that labels/titles for icons, menus and symbols accurately describe their content.	4.3.2	Gelderblom [2008]
9.2	Component labels/titles should not mislead users into accessing content they would otherwise not be interested in.	4.3.2	Gelderblom [2008]
10	**The Digital Doorway should support multiple ways of interaction with the user**		
10.1	The Digital Doorway should not impose unnecessary constraints on the user input method.	4.3.1.2	Dix et al. [2004]
10.2	Where user input can be provided via the keyboard and onscreen keys, the user should be allowed to provide input through either method.	4.3.1.2	Dix et al. [2004]
10.3	The Digital Doorway should support multiple output methods.	4.3.1.2	Dix et al. [2004]

7.3.2 Category 2: Form Usability Heuristics

The form usability heuristics are primarily based on the principles and guidelines discussed in section 7.2.1. There are five high level heuristics and twelve sub-heuristics in the category. The description of the form usability heuristics for DD evaluation is provided in Table 7.2.

Table 7.2: Form usability heuristics for Digital Doorway evaluation

H_No	Heuristics	Applicable sections	Reference
1	**Provide support for easy navigation around form elements**		
1.1	Provide visible cue by positioning the cursor in the first field at start of the form.	4.4.1.1 5.5.3 5.5.4	Mayhew [1992] Section 508 [2000] IBM [2009]
1.2	Cursor movement should follow the order in which form elements are organized.	4.4.1.1 5.5.3	Mayhew [1992] Section 508 [2000]

		5.5.4	IBM [2009]
1.3	Users should be able to edit data fields by moving the cursor backward and forward, rather than having to retype the whole field.	4.4.1.1	Mayhew [1992]
2	**Organize form elements in a logical way**		
2.1	Ensure that related items are grouped together to aid readability.	4.4.1.1	Mayhew [1992]
2.2	Provide visual reinforcement for element groups through efficient use of white spaces and borders.	4.4.1.1	Mayhew [1992]
3	**Provide adequate information to enable successful completion of form**		
3.1	Ensure that required information is clearly specified.	4.4.1.1	Mayhew [1992]
3.2	Designate required fields in standard and consistent ways, taking into account the user's age and knowledge.	4.4.1.1	Mayhew [1992]
3.3	Give feedback for missing data fields in clear and unambiguous terms, taking into account the user's age and knowledge.	4.4.1.1	Mayhew [1992]
3.4	When input errors are detected, the cursor should be positioned in the error field with the field highlighted to attract the user's attention.	4.4.1.1	Mayhew [1992]
4	**Ensure that data entry fields are associated with appropriate captions/labels**		
4.1	Give meaningful names to field captions/labels, taking into account the user's age and experience.	4.3.1.1 4.3.2 4.4.1.1	Dix et al. [2004] Gelderblom [2008] Mayhew [1992]
4.2	Ensure that captions/labels are distinct from data entry fields.	4.4.1.1	Mayhew [1992]
5	**Ensure that data entry field length is sufficient for the size of required data**		
5.1	Clearly specify the limit for data that has minimum or maximum allowable length.	4.4.1.1	Mayhew [1992]

7.3.3 Category 3: Heuristics to Support Direct Accessibility

Category 3 heuristics, aimed at evaluating the level of direct accessibility support in the DD are derived from the principles and guidelines discussed in section 7.2.1. The heuristics, shown in Table 7.3, consists of five high level heuristics and fifteen sub-heuristics.

Table 7.3: Direct accessibility heuristics for Digital Doorway evaluation

H_No	Heuristics	Applicable sections	Reference
1	**Provide information that is perceptible to users with different ability**		
1.1	Font size of instructions should be large enough to enable easy perception by users with low vision.	5.5.1.1 5.5.4.	Story et al. [1998] IBM [2009]
1.2	Information should be accessible without undue physical efforts.	5.5.1.1	Story et al. [1998]
1.3	Important information should be clearly distinguishable from other peripheral contents.	5.5.1.1	Story et al. [1998]

1.4	Provide audio equivalent of instructions and information to afford access by users who cannot read.	5.5.1.1 5.5.2 5.5.4.	Story et al. [1998] WCAG 1.0 [1999] IBM [2009]
1.5	Provide quality speech output that enable users to hear and comprehend their meanings.	5.5.4	IBM [2009]
2	**Ensure that feedback is accessible to users regardless of their ability**		
2.1	Provide feedback using multiple modes to facilitate access and comprehension.	4.3.2 5.5.1.1 5.5.2 5.5.3 5.5.4	Gelderblom [2008] Story et al. [1998] WCAG 1.0 [1999] Section 508 [2000] IBM [2009]
2.2	Text equivalent of graphic or audio information should convey the same message.	5.5.2 5.5.4	WCAG 1.0 [1999] IBM [2009]
3	**Do not rely on colour alone to code and distinguish**		
3.1	Ensure that colour alone is not used to represent important information.	5.5.2 5.5.3 5.5.4	WCAG 1.0 [1999] Section 508 [2000] IBM [2009]
3.2	Ensure that background and text colours contrast well with each other.	5.5.2 5.5.3 5.5.4	WCAG 1.0 [1999] Section 508 [2000] IBM [2009]
4	**Allow complete and efficient usage of the keyboard**		
4.1	Allow keyboard navigation for operations/tasks that do not essentially require use of the mouse.	5.5.2 5.5.3 5.5.4	WCAG 1.0 [1999] Section 508 [2000] IBM [2009]
4.2	Ensure that menus and buttons are accessible using the keyboard.	5.5.2 5.5.3 5.5.4	WCAG 1.0 [1999] Section 508 [2000] IBM [2009]
5	**Allow user control of audio-visual information**		
5.1	Avoid automatic progression from one screen to the next for audio-visual information. Users should explicitly select forward/backward progression.	5.5.2 5.5.4	WCAG 1.0 [1999] IBM [2009]
5.2	Provide controls that enable users to pause, continue, or repeat audio-visual information.	5.5.2 5.5.4	WCAG 1.0 [1999] IBM [2009]
5.3	Users should be able to adjust the volume of audio information	5.5.2 5.5.4	WCAG 1.0 [1999] IBM [2009]
5.4	Equivalent audio information should be synchronized with the text alternatives.	5.5.4	IBM [2009]

7.3.4 Category 4: Educational Game Usability Heuristics

Category 4 heuristics are aimed at evaluating the usability of *What-What Mzansi*, *OpenSpell* and *Themba's Journey*. The heuristics (Table 7.4) are derived the guidelines discussed in section 7.2.1. There are seven high level heuristics and 21 sub-heuristics in the category.

Table 7.4: Educational game usability heuristics for Digital Doorway evaluation

H_No	Heuristics	Applicable sections	Reference
1	**Ensure that the goals, aims and objectives of the game are explicitly specified**		
1.1	Games should have clear goals and objectives.	6.5.2 6.5.3 6.5.4	Alessi and Trollip [2001] Malone [1980, 1981] Desurvire et al [2004]
1.2	Ensure that learners can easily determine whether they are getting closer to the goal.	6.5.2	Alessi and Trollip [2001]
2	**Make learners aware of the rules of the game**		
2.1	Provide an easily accessible instruction on how to play the game.	6.5.2 6.5.4	Alessi and Trollip [2001] Desurvire et al [2004]
2.2	Permissible actions should be clearly specified.	6.5.2 6.5.4	Alessi and Trollip [2001] Desurvire et al [2004]
2.3	Clearly specify constraints and restrictions governing the game.	6.5.2 6.5.4	Alessi and Trollip [2001] Desurvire et al [2004]
3	**Provide appropriate level of learner control**		
3.1	Learners should be able to adjust the game's level of difficulty.	6.5.2 6.5.3 6.5.4	Alessi and Trollip [2001] Malone [1980, 1981] Desurvire et al [2004]
3.2	The application should be able to adjust the level of difficulty based on the learner's performance.	6.5.2 6.5.3 6.5.4	Alessi and Trollip [2001] Malone [1980, 1981] Desurvire et al [2004]
3.3	Whenever appropriate, give learners the option of returning to where they left off when the program is temporarily exited.	6.5.2	Alessi and Trollip [2001]
4	**Provide appropriate level of challenge for learner motivation and self-esteem**		
4.1	Ensure constant challenge through adjustable difficulty level.	4.3.2 4.3.4.2 6.5.2 6.5.3 6.5.4	Gelderblom [2008] Preece et al. [2007] Alessi and Trollip [2001] Malone [1980, 1981] Desurvire et al [2004]
4.2	Performance feedback should not be given using negative or sarcastic statements.	6.5.2 6.5.3	Alessi and Trollip [2001] Malone [1980, 1981]
4.3	Provide constructive and corrective feedback that will enable the learner to learn from mistakes and improve future performance.	6.5.2 6.5.3	Alessi and Trollip [2001] Malone [1980, 1981]
5	**Provide support for the correction of cognitive mistakes**		
5.1	Learners should have more than one opportunity to provide answers.	4.3.2	Gelderblom [2008]
5.2	Appropriate hints should be provided for the correction of cognitive mistakes.	4.3.2 6.5.4	Gelderblom [2008] Desurvire et al [2004]
6	**Provide adequate control mechanisms to support easy navigation**		
6.1	Provide learner control options for forward progression to facilitate skipping a section and	6.5.2	Alessi and Trollip [2001]

	backward progression, which enables the review of a previous section.	6.5.4	Desurvire et al. [2004]
6.2	All control mechanisms should be visible and easily accessible.	6.5.2	Alessi and Trollip [2001]
6.3	All control mechanisms should be easy to use without requiring undue physical efforts.	5.5.1.1 6.5.2	Story et al. [1998] Alessi and Trollip [2001]
6.4	Provide clear exit mechanism to allow learners leave the game at any stage.	6.5.2	Alessi and Trollip [2001]
6.5	Request to terminate the program should be confirmed by the learner to avoid unintentional exit.	6.5.2	Alessi and Trollip [2001]
7	**Recognize and respect users' socio-cultural and language diversity**		
7.1	Game should be accessible in different languages.	4.3.2 6.5.1	Gelderblom [2008] Shelley [2001]
7.2	Game content should not be biased against specific cultural or gender groups.	4.3.2 6.5.1	Gelderblom [2008] Shelley [2001]
7.3	Game activities should be embedded in scenarios that learners can relate to.	4.3.2 6.5.1	Gelderblom [2008] Shelley [2001]

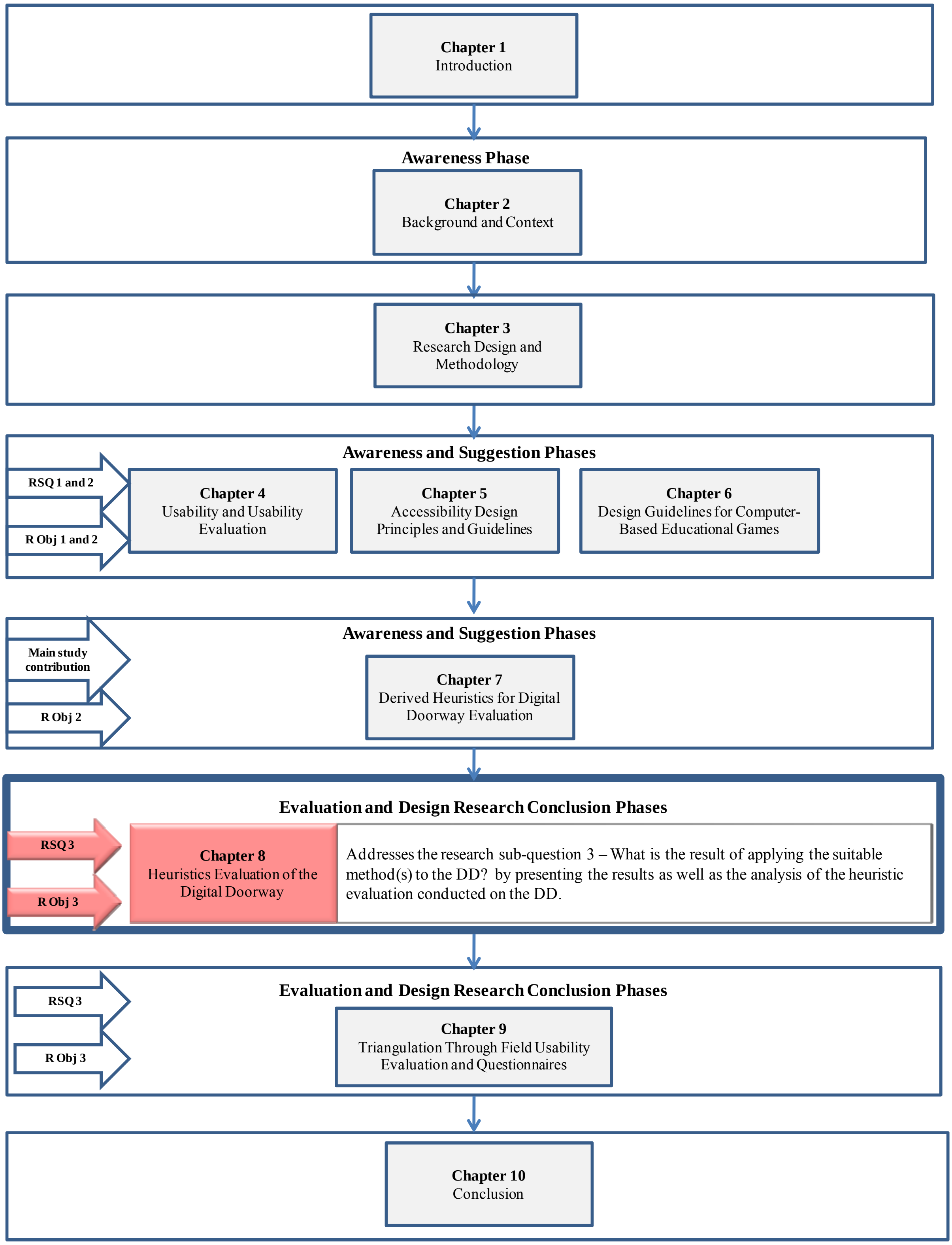
7.4 CONCLUSION

In this chapter, I have presented the main contribution of this study, that is, the set of multi-category heuristics specifically developed for evaluating a selection of interfaces and applications on the DD. The evaluation heuristics were designed to address both the general usability and direct accessibility support provided in the DD. Furthermore, heuristics focussing on the usability of education game applications were included in the heuristic set. In presenting the multi-category heuristics, cross-references were made to specific sections of the dissertation where the principles and guidelines that form the basis for the heuristics were discussed and references for the author(s) who proposed the principles and guidelines.

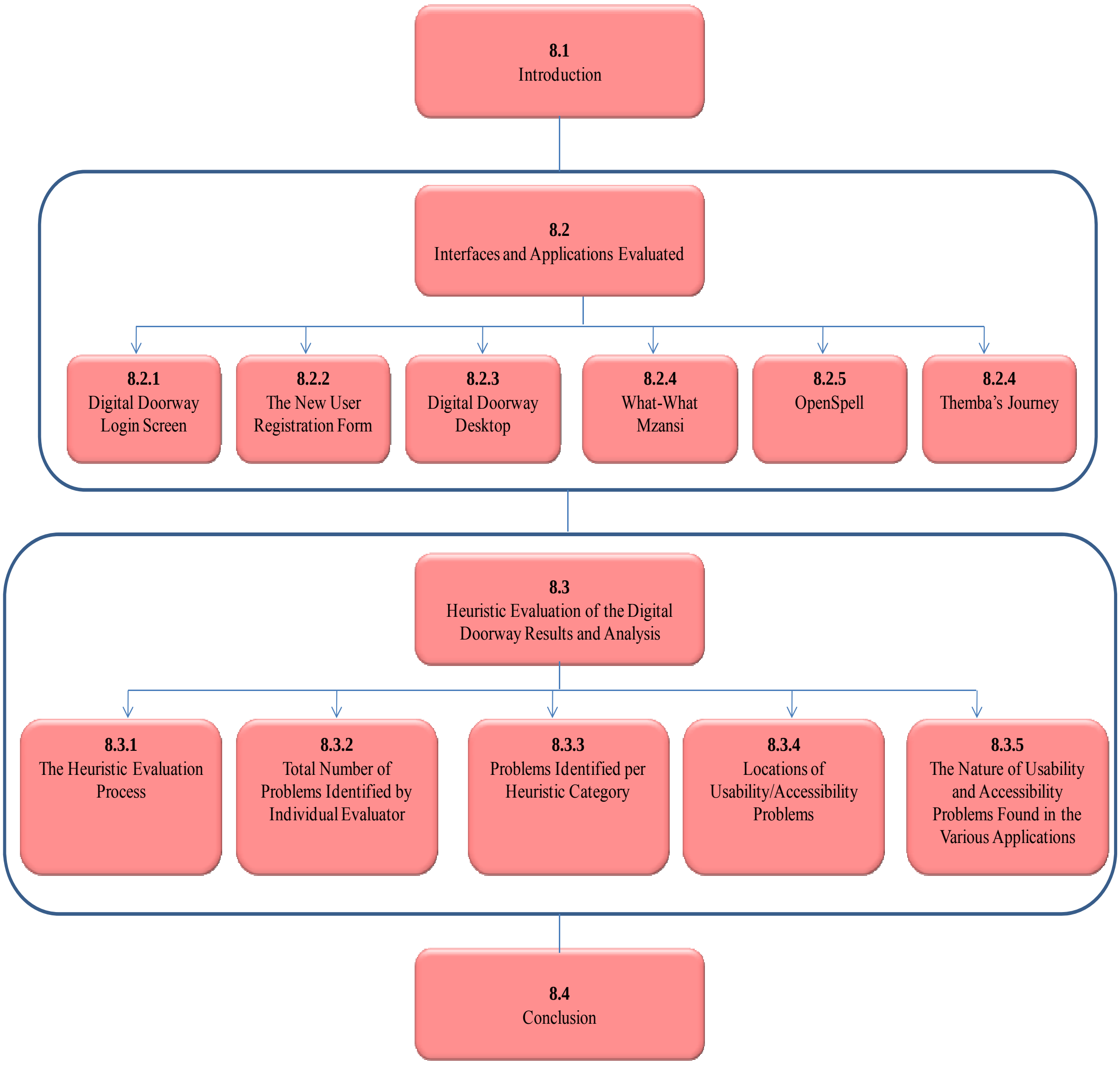
In Chapter 8, I will illustrate the applicability of the heuristics by presenting the result of the formal heuristic evaluation of the selection of interfaces and applications on the DD while Chapter 9 will triangulate the outcome of the evaluation through a field study. The purpose of this is to determine the effectiveness of the heuristics using the correctness, coverage and terminology criteria proposed by Sim et al. [2009].

CHAPTER 8: HEURISTIC EVALUATION OF THE DIGITAL DOORWAY

The stage of Chapter 8 in the dissertation



Map of Chapter 8



8.1 INTRODUCTION

This chapter partly maps onto the evaluation activities and the conclusion phase of the outer cycle of the design research process. I provide part of the answer to my research sub-question 3: What is the result of applying the suitable method(s) to the DD? starting with the description of the interfaces and applications evaluated in section 8.2. In section 8.3, I present the detailed description of the formal heuristic evaluation process. This is followed by the discussion of the results obtained from the heuristic evaluation where the multi-category heuristics presented in section 7.3 were utilized. I present the analyses of the results along with the results. Analysis was done based on the number of problems identified by individual evaluator, the category of the heuristics violated, the interfaces and applications where problems are located, and the nature of the problems identified in the various applications. Section 8.4 concludes the chapter.

8.2 INTERFACES AND APPLICATIONS EVALUATED

As mentioned in section 2.4.3, the DD provides a large number of applications and resources for user access. The sheer size of available materials made it infeasible to evaluate all of them. Furthermore, the majority of these resources are open source and third-party applications, and evaluating these third-party resources may not necessarily lead to improvement since the DD project team has no control over their design.

To ensure that the results from this study positively influence future development efforts by the DD project team, I decided to evaluate interfaces and applications that were developed in-house for the DD. The interfaces and applications evaluated are: the login screen, the registration form for creating a new user account, the main desktop, and three educational games, *What-What Mzansi*, *OpenSpell* and *Themba's Journey*. The following sub-sections describe these interfaces and applications.

8.2.1 Digital Doorway Login Screen

The login screen, shown in Figure 8.1, is the first interface between the user and the DD. Users access DD content by logging in as a guest or registered user. The main language of instruction/information on the login screen is English. However, equivalent information is provided in four other South African languages, namely, IsiXhosa, Afrikaans, Sotho, and Venda.

A guest user can simply access content by typing 'dd1' in the username textbox. Alternatively, typing

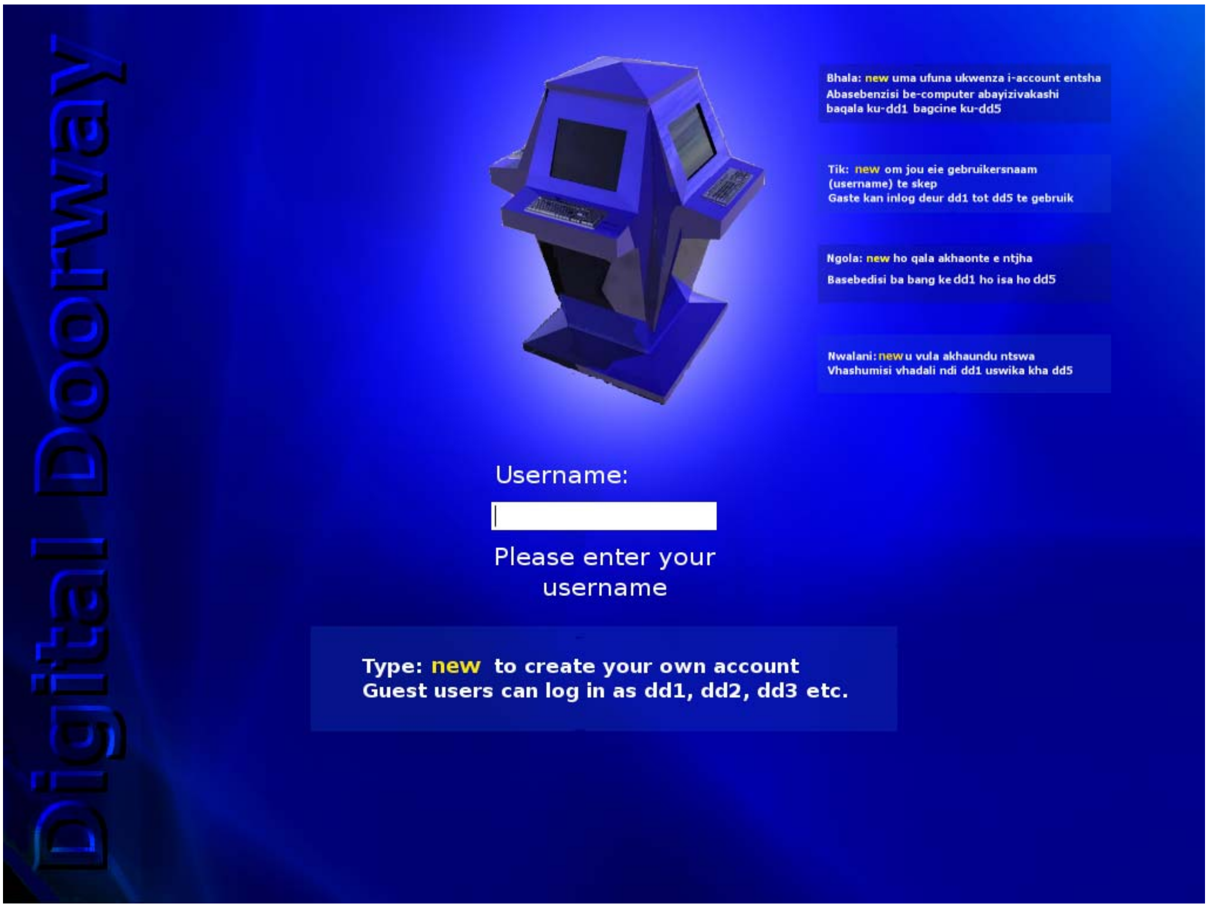


Figure 8.1: Digital Doorway login screen

'new' enables creating user accounts by activating the registration form. Although not specifically stated on the login screen, one benefit of using the DD as a registered user involves Internet access through the Global Packet Radio Service (GPRS) technology, which is currently being piloted.

DD users can access the Internet by purchasing a standard prepaid mobile phone recharge vouchers. If a registered DD user does not use all the Internet time available on a voucher, s/he can use the remaining credit balance at a later stage because the access detail is captured and stored. A guest user risks having his/her remaining credit used by another guest who log in with the same guest 'name'.

Users interact with the login screen using the metal touchpad and keyboard. First, the user positions the pointer in the username textbox by pressing the left 'mouse click' key on top of the metal touchpad (see Figure 2.5). Thereafter, the user provides the username by pressing the applicable alpha-numeric keys on the metal keyboard. The end of the username input is signalled by pressing the 'Enter' key.

If the name provided is that of a guest, for example 'dd1', the user is logged in and presented with the main desktop (discussed in section 8.2.3). If the user is previously registered, the username textbox is replaced with a password textbox. The user password is provided in a similar way to the username. If both the username and password are correct, the user is logged in and presented with the main desktop. If the username and/or password is/are incorrect, the user is presented with the username textbox and the password textbox in succession until the correct username and password are provided.

8.2.2 The New User Registration Form

DD users may choose to create a user account by completing a simple electronic form (Figure 8.2). Items on the form are organized into two main groups: the "Personal Details" and "User Details" groups. Within the personal details group, user information such as name, age and gender can be provided. A user-selected username and password are chosen within the user detail group. The form also provides users with hints on the type of data expected at certain fields, for example the password field.

The main user interaction devices with the registration form are the metal touchpad and keyboard. To begin input, users must first position the insertion point in a data field using the touchpad and press the left 'mouse click' key. The user fills in the full name and age fields in the personal details group, and user details fields by pressing the applicable alpha-numeric keys on the metal keyboard. Users select gender by positioning the pointer over one of the radio buttons and pressing the left 'mouse click' key. The applicable home language and the preferred language are selected from the choice of languages provided in the drop-down lists in a similar way.

Although this is not explicitly stated on the form, the form requires all data fields to be filled but data fields can be entered in any order. At the completion of all the data fields, the user registration information is stored following a 'click' on the <Register User> button. A click on the <Cancel> button will close the form without providing any warning to the user.

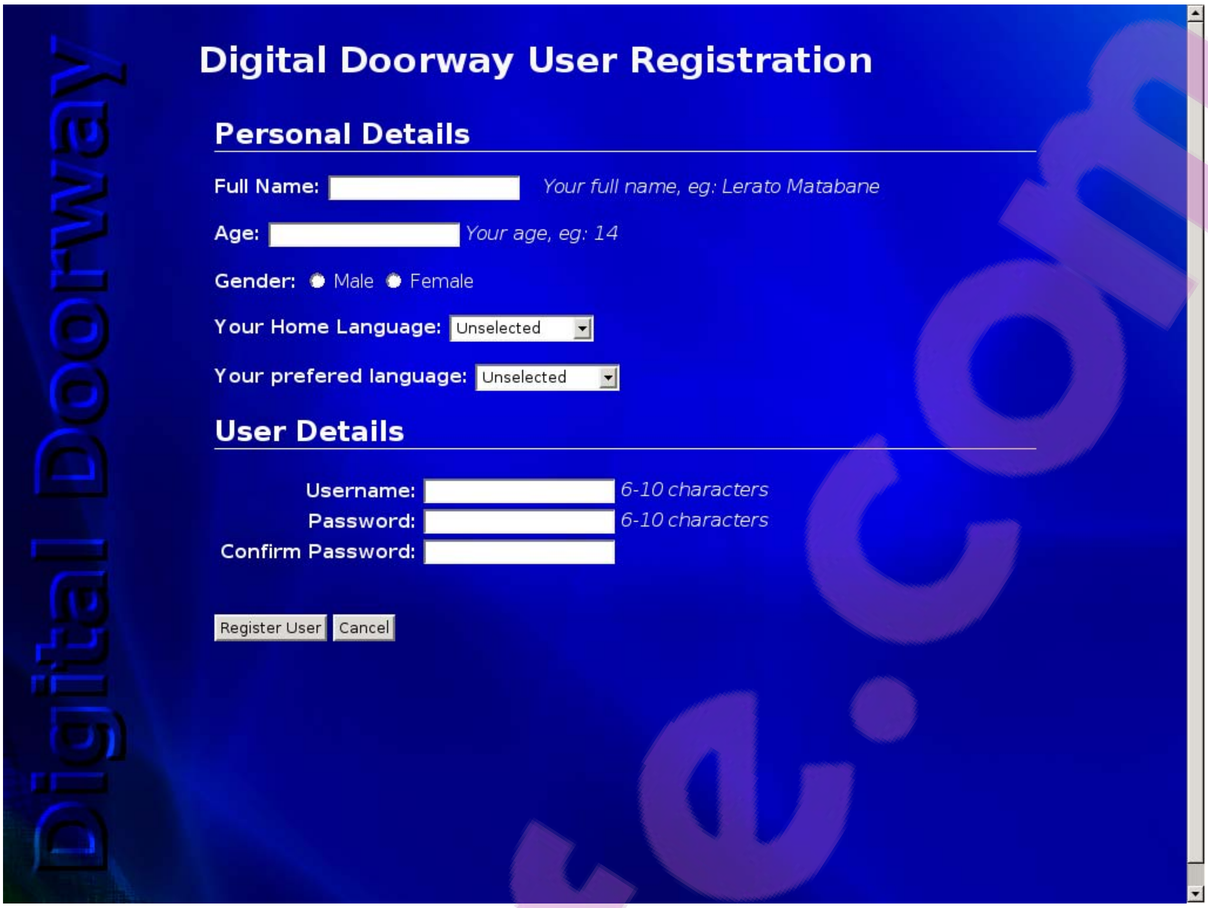


Figure 8.2: Digital Doorway user registration form

8.2.3 Digital Doorway Desktop

Following a successful login, applications and content on the DD can either be opened by clicking on icons on the desktop or by selecting from the two menu options: "Programs" and "Resources". The desktop also provides global volume controls in two ways. The user can click on a 'volume control' icon, shown as the green slider bar on the extreme right of the taskbar in Figure 8.3. Alternatively, the user can access a more advanced volume control dialogue window by clicking on the icon just before the green slider bar. Users can log out of the system by either clicking on an 'exit' button (designated by a right pointing arrow ⇒) or from the advanced system menu (located on the taskbar).

Users interact primarily with the main desktop using the metal touchpad and the 'mouse click' keys. Applications such as the educational game *What-What Mzansi*, can be launched by positioning the pointer over the new_content folder, for example, and pressing the left or right 'mouse click' key.

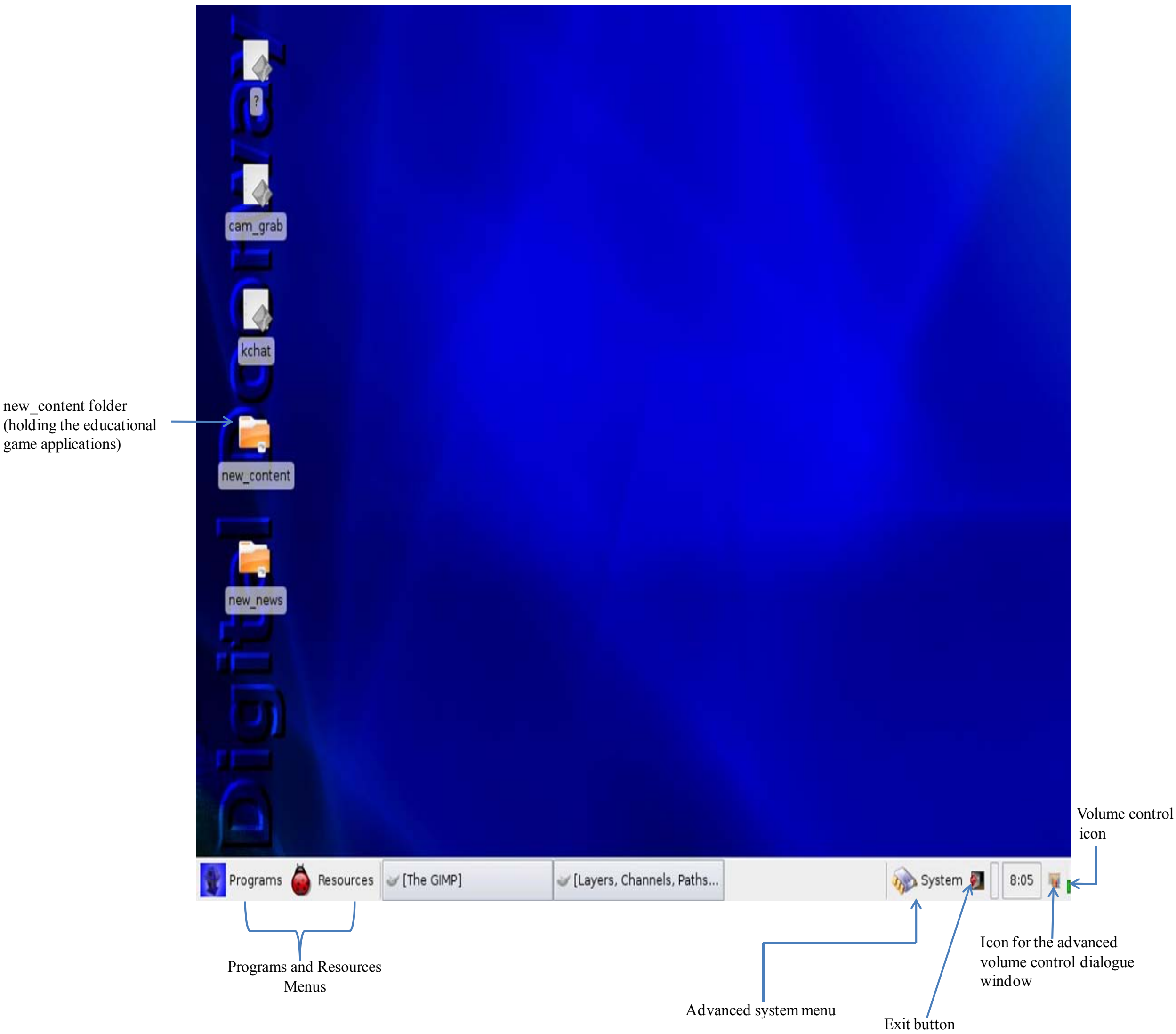


Figure 8.3: Digital Doorway desktop

8.2.4 What-What Mzansi

What-What Mzansi is a general knowledge quiz game in the form of yes/no questions. The interface (Figure 8.4) provides three menu items: play, hi-Scores and about:

1. *Play:* When the user selects the play menu option the interface asks the questions. The player selects from two levels of difficulty: "Easy" and "Advanced" and then the program character welcomes the player and reads out the questions. Each session lasts sixty seconds. The score for each question can range from two to ten, depending on how fast the player is able to answer it. As shown at the bottom right hand side of Figure 8.5, which is activated when the play option is selected, the maximum score for the current question is two marks, if the answer is correct. If the answer is wrong, the score value is deducted from the running total score. At the end of the session, all the questions are presented again with tick marks or crosses over the questions to show whether or not the player answered them correctly. If the player performed well, the

program character informs him/her of his/her readiness to move to the next level. If performance is poor, s/he can choose to repeat the current level or elect to exit the application.

2. *Hi-Scores:* This menu lists the scores of the top-ten registered users.
3. *About***:** This menu option mainly provides information on the DD project and its achievements.

The '?' icon, on the top right corner, provides context-specific instructions when a user clicks on it, while the 'X' icon closes the application.

Users interact with *What-What Mzansi* using the metal touchpad, the keyboard and the 'mouse click' keys. Questions can be answered in one of two ways:

1. By positioning the pointer over the 'yes' or 'no' pot (Figure 8.5) and then pressing the left 'mouse click' key.
2. By pressing the left arrow key on the keyboard to answer 'yes' or the right arrow key to answer 'no'.

Figure 8.4: Interface of What-What Mzansi

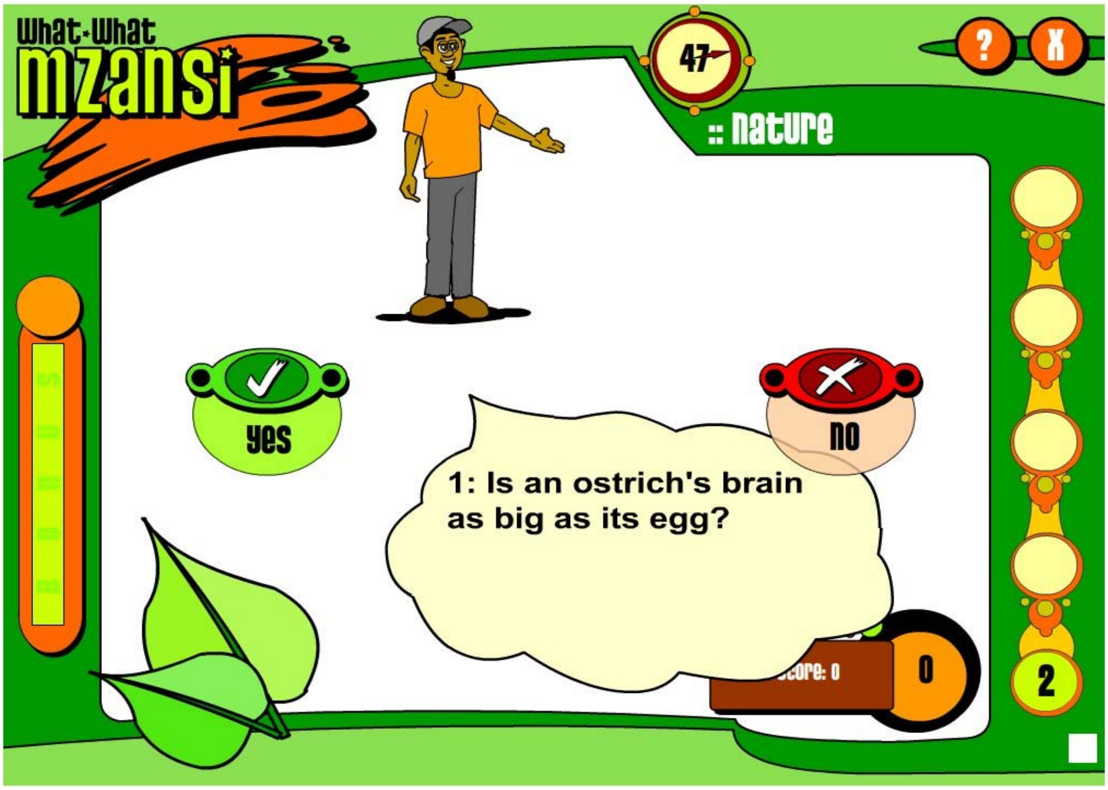


Figure 8.5: Screenshot from What-What Mzansi

8.2.5 OpenSpell

OpenSpell is an educational spelling game. Available in all the eleven official South African languages, it provides three levels of difficulty, designated with *, **, ***, to represent the simplest, intermediate, and the highest level of difficulty respectively.

The interface (shown in Figure 8.6) presents three game options: say, guess and spell. The player uses onscreen alphabetical keys when completing spelling and guessing exercises. The three control buttons, <repeat>, <erase> and <enter>, are active only when the spell option is selected. The primary interaction devices with *OpenSpell* are the metal touchpad and the left ‘mouse click’ key. The functionalities of the three game options are:

1. *Say:* This option brings up pictures of the words to be spelt one after the other. A voice in the chosen language speaks out each letter of specific word and the pronunciation of that word.
2. *Guess:* This functionality is based on the hangman word guessing game and presents the letters of the word as a series of dashes. The player selects letters of the alphabet from the onscreen key by positioning the pointer over it and pressing the left ‘mouse click’ key above the metal keyboard. If the chosen letter appears in the word, it is slotted in the appropriate space(s). With each incorrect guess, a bird is perched on a tree branch accompanied by a sound effect of the branch breaking off the tree. This process is continued until the player guesses the correct word or the tree branch breaks.

3. *Spell:* This option prompts the player to spell a word that is read out and described in a picture. The player positions the pointer over the onscreen alphabetic keys and presses the left ‘mouse click’ key on the metal keyboard. The end of word spelling is signalled by clicking the <enter> button. Spell permits the player to attempt a word a second time if the first attempt is wrong, then the application provides the correct answer. Players can correct an erroneous input by clicking the <erase> button, erasing the word letter by letter.

The user selects from three levels of difficulty and as the performance and confidence improves, s/he can then select a more challenging level. Availability of the program in all the official South African languages reflects developers’ awareness of diversity in the country. This also supports accessibility.

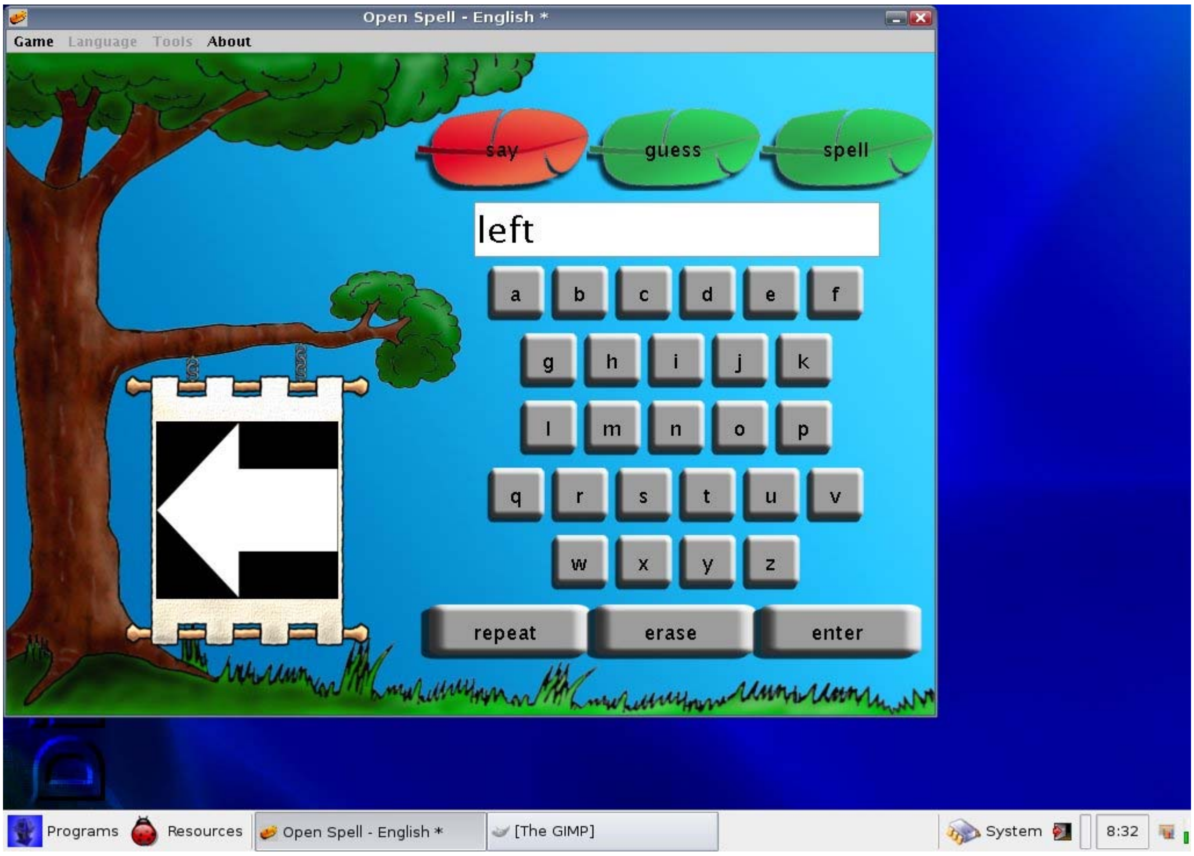


Figure 8.6: Interface of OpenSpell (executing in <say> mode)

8.2.6 Themba’s Journey

Themba’s Journey is a life skills program, available in IsiXhosa and English, and tells the story of Themba, who makes a journey from his village to the city. The main interface (shown in Figure 8.7) provides three menu items: Play, Help and Exit:

1. *Play* narrates the main story where at strategic places the user has to make decisions. Users make decisions on a course of action on Themba’s behalf when he reaches crossroads such as whether to take drugs or not. Each direction can lead to positive or negative consequences. The default

language is IsiXhosa, available in text and spoken words. An English equivalent is only available in text, which is revealed when the user hovers the pointer over the speech bubbles.

2. *Help* provides navigation and game instructions.
3. *Exit* closes the program.

Themba's journey is not a typical educational game but rather one that follows a story line. It relies on users' ability to deduce the causal relationships between the path taken at crossroads and the lessons that are being passed across. The context in which the narration is set mirrors typical city life in South Africa, which many users can relate to.

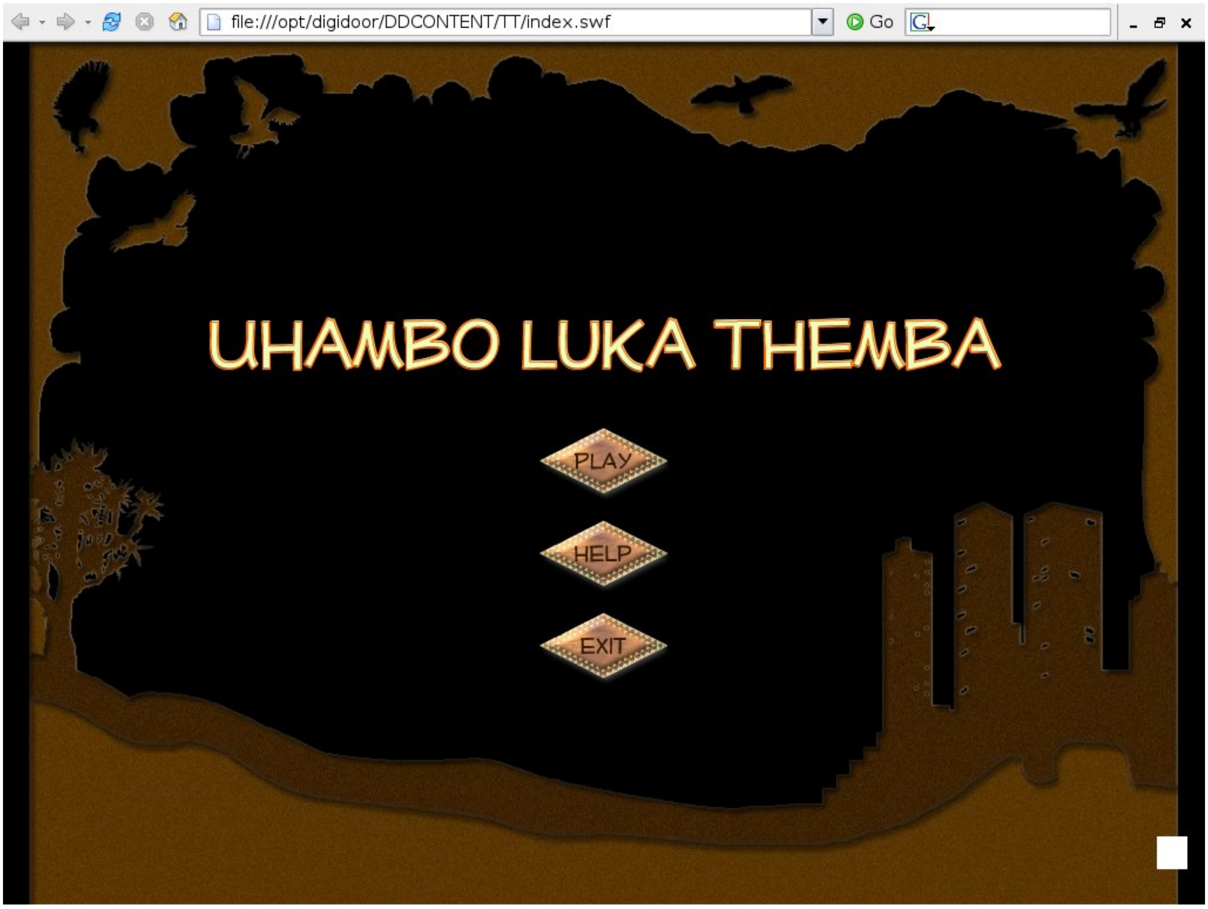
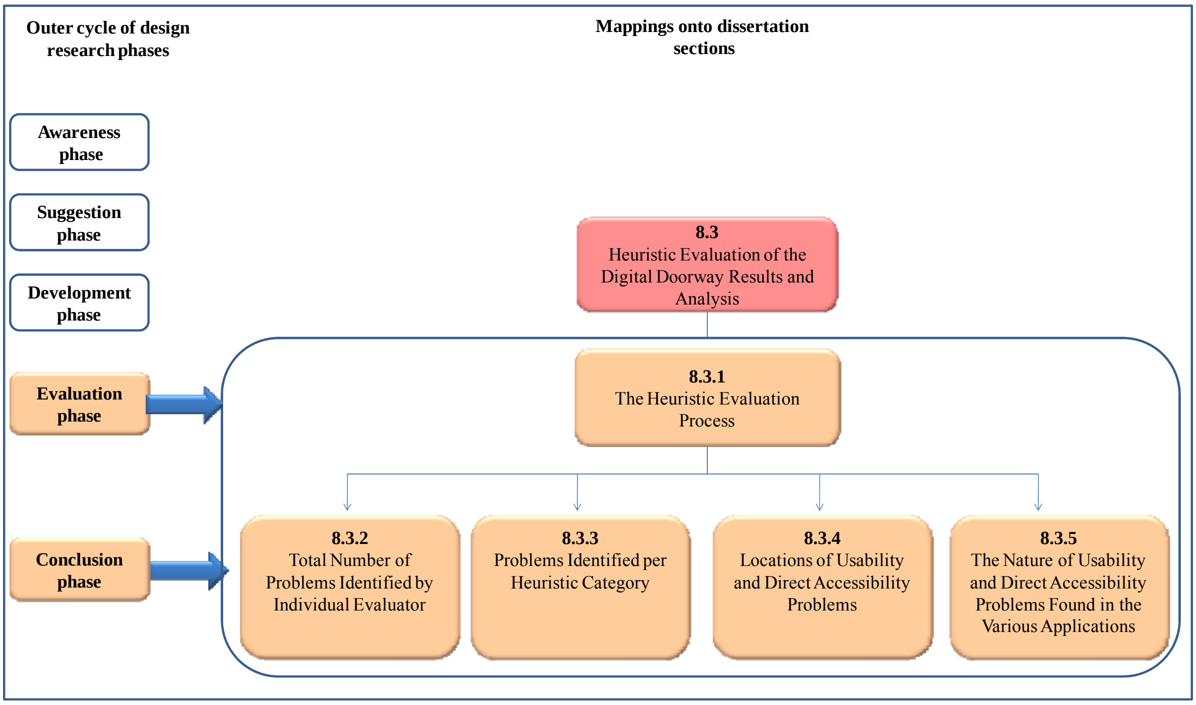


Figure 8.7: Interface of Themba's Journey

8.3 HEURISTIC EVALUATION OF THE DIGITAL DOORWAY RESULTS AND ANALYSIS

Mapping of section 8.3 to the design research phases



This section maps onto the evaluation and conclusion phases of the outer cycle of the design research and provides part of the answer to my research sub-question 3: What is the result of applying the suitable method(s) to the DD?

In Chapter 7, I presented the set of multi-category heuristics developed for performing a heuristic evaluation on a selection of interfaces and applications installed on the DD, discussed in section 8.2. Section 8.3.1 describes the process I followed in performing the heuristic evaluation. Sections 8.3.2 to 8.3.5 present the results obtained from the heuristic evaluation and the analysis of the results.

8.3.1 The Heuristic Evaluation Process

A team of five experts independently evaluated the DD using the developed multi-category heuristics (section 7.2.2.3). For logistical reasons, evaluators came physically to the DD laboratory located at the CSIR for the evaluation. The five evaluation sessions were conducted over a period of three months. An initial ten-week period, between 15 September 2009 and 30 November 2009, was scheduled for each evaluator to find a convenient time to conduct the evaluation. At the end of this period, only three evaluators were able to evaluate the DD due to work commitments. It was infeasible to schedule evaluation sessions during the month of December due to seasonal festivities. Thus the remaining evaluators conducted the evaluation in January and February 2010.

Two factors were considered when selecting evaluators for the formal heuristic evaluation of the DD. Firstly, the relationship between potential evaluators' expertise and effective evaluation results meant the potential evaluators approached for participation were those with usability and/or accessibility evaluation experience. Secondly, in line with Nielsen's [1994b] recommendation for between three and five expert evaluators, five expert evaluators participated in the heuristic evaluation of the DD. As their profiles show (Table 8.1), three evaluators had experience in usability evaluations while the other two had expertise in usability and accessibility evaluation.

An information document (Appendix E) was mailed to all evaluators prior to the evaluation to ensure uniformity in briefing and allow evaluators sufficient familiarization time. The document overviewed the interfaces and applications to be evaluated; typical user profiles; procedures to be followed during evaluation; and, the multi-category heuristics (see section 7.3).

Using Nielsen's [1994b] procedure for conducting heuristic evaluation, with some modifications, evaluators undertook a two-pass session. The first aimed at getting a feel of how the applications work; and the second aimed at the evaluation. No time limit was set but each evaluation session lasted approximately two hours and one evaluator spent nearly two hours and forty minutes.

Because of the nature of the system, the evaluators had to stand in front of the DD to open, interact with and evaluate the interfaces and applications selected for evaluation with short breaks in-between applications. This was not conducive for evaluators to evaluate and record notes of their findings simultaneously. All five evaluators accepted my offer to act as scribe in order to ease the evaluation tasks.

At the start of each evaluation session, I discussed the heuristics with the evaluators. Then, each evaluator traversed the interfaces and applications one after the other by performing typical user tasks and described usability/accessibility problems relating to the specific interface or application to me. After the session, I compiled an evaluation report which I mailed to the relevant evaluator the day after the evaluation (only one evaluator received the report two days after the evaluation). Then, the evaluator verified the report to ascertain whether it was a true reflection of the evaluation. In some cases, the verification process resulted in modifications to the report by evaluators. Rather than

Table 8.1: Expert evaluators' profile

Evaluator	Qualification	Position/Job Title	Duties/Roles at Institution
1	MSc (Information Systems)	Senior Lecturer	Tuition; Postgraduate supervision
2	MSc, Certified Usability Analyst	Usability manager and Researcher	Usability analyst/consultant; Usability and Eye tracking research
3	MSc (Computer Science)	Researcher	Research in the field of optimization; Organization and conduct of usability and accessibility evaluations
4	PhD, MSc, Med	Full Professor	Tuition; Postgraduate supervision; Research; Management and Leadership
5	M(Eng): Technology Management	Researcher	Research in the field of voice user interface; Conduct of usability and accessibility evaluations

expecting evaluators to produce a full report from scratch this approach made judicious use of evaluators' time and focused evaluators' attention on the evaluation process.

Following the five heuristic evaluation sessions, the evaluators identified a large number of usability and direct accessibility-related problems. I aggregated these into a list of 71 usability and direct accessibility problems by consolidating problems that were of similar nature. For example, I recorded a single problem for the absence of feedback for an incorrect username and/or password on the login screen, detected by four of the five evaluators. Likewise, I recorded a single problem relating to the low contrast between the background of the main desktop and the labels for icons on the desktop, as well as between the desktop and the word 'Digital Doorway' written across the desktop.

Similar problems affecting different interfaces and applications were listed as separate problems to ensure that each received the necessary attention. For example, the two problems relating to small font size on the login screen and in *OpenSpell* were listed as two separate problems.

The calculation of severity ratings (see section 4.2.3.1.6) can guide decisions about resource allocation in correcting usability problems [Nielsen, 1994b]. However, this study did not insist on severity ratings as this might have resulted in evaluator subjectivity because some evaluators would have had to remember problems over prolonged periods. As mentioned above, scheduling factors meant the five evaluation sessions took place over a period of three months; thus the time elapsed between some evaluation sessions and producing the aggregated list of problems was months. Further, evaluators could not easily access the DD again at their convenience.

Following the formal heuristic evaluations, two of the problems identified by evaluators could not be matched to any of the multi-category heuristics (see section 7.2.2.3); thus additional heuristics were generated to cover these problems.

The remainder of this section presents an analysis of the results of the formal heuristic evaluation.

8.3.2 Total Number of Problems Identified by Individual Evaluator

As shown in Table 8.2, Evaluator 1 and 4 identified 32 problems; Evaluator 3 identified 33; Evaluator 2 identified 34 and Evaluator 5 39 problems. It is interesting to note that all the evaluators identified a similar narrow proportion of problems (between 45 and 55 percent), as clearly illustrated in Figure 8.8, but that none of the evaluators was able to identify a substantial proportion of the total number of problems. It is therefore valuable to examine the variations in the set of problems identified by the various evaluators.

Table 8.2: Number of problems identified by evaluators

Evaluator	Number of problems identified	Problem percentages*
1	32	45%
2	34	48%
3	33	46%
4	32	45%
5	39	55%
*Problem percentages are relative to the aggregated 71 usability/accessibility problems.		

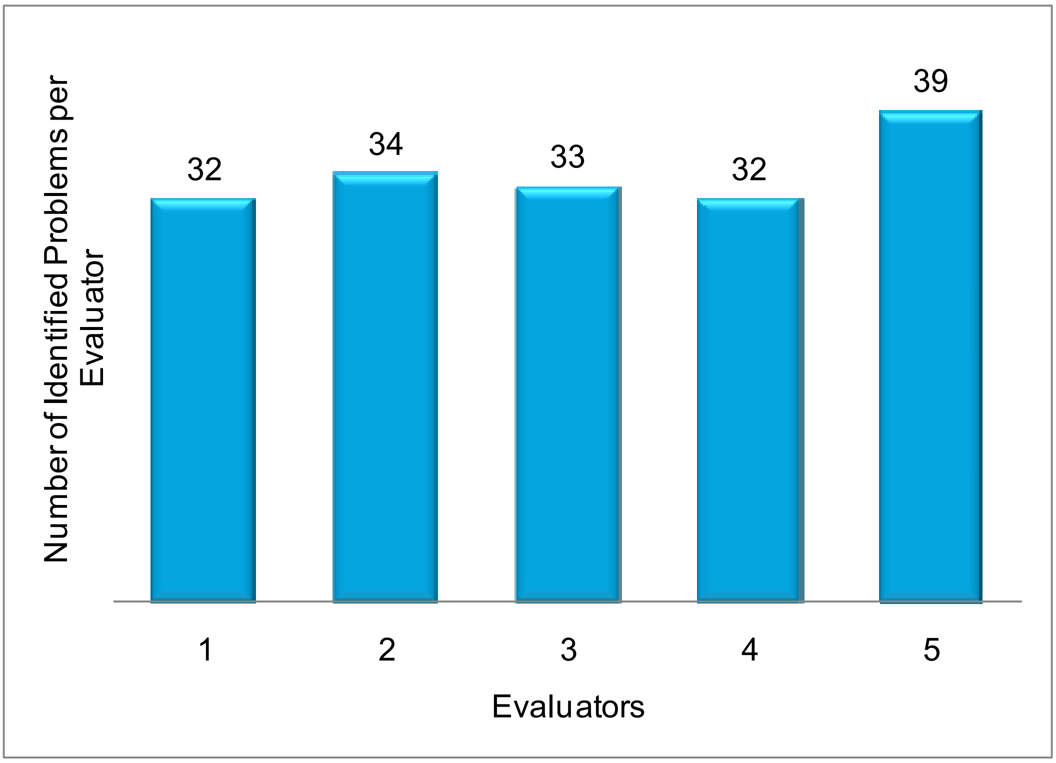


Figure 8.8: Total number of problems identified by each evaluator

Analysis of the variation in the set of usability and direct accessibility-related problems identified by the five evaluators considered the aggregated list of 71. The first considered the degree of agreement among evaluators in identifying a specific problem. As shown in Table 8.3, problems identified by single evaluators accounted for 34 of the aggregated list of problems while ten were recognized by two evaluators. Seven problems were identified by three evaluators and a further seven were identified by four evaluators. All five evaluators agreed that thirteen of the aggregated list of problems were indeed problems.

Further analysis of problem detection variation between evaluators explored the specific interface and application in which the problem occurred. As shown in Table 8.4, only one of the eight problems recorded for the login screen was detected by all five evaluators and half of the problems detected were by single evaluators. None of the four problems recorded for the registration form was detected

Table 8.3: Number of problems detected by single and multiple evaluators

Recognized by	No of problems	Problem percentage
One evaluator	34	47.88%
Two evaluators	10	14.08%
Three evaluators	7	9.86%
Four evaluators	7	9.85%
All evaluators	13	18.30%
At least three evaluators	27	38.02%

by all five evaluators and two were identified by single evaluators. Five of the thirteen problems detected for the main desktop were recognized by single evaluators and two problems were recorded by all the evaluators.

The highest disagreement between evaluators in identifying problems occurred for *What-What Mzansi* and *Themba's Journey:* nine problems were identified by single evaluators in both applications. The best agreement between evaluators in identifying problems occurred for *OpenSpell,* where five of the fourteen problems in the application were detected by all evaluators, the highest for all the interfaces and applications.

Table 8.4: Variations in number of problems identified per interface/application

	Interfaces/Applications					
Recognized by	**Login Screen**	**Registration form**	**Main desktop**	**What-What Mzansi**	**OpenSpell**	**Themba's Journey**
One evaluator	4	2	5	9	5	9
Two evaluators	1	1	2	3	2	1
Three evaluators	1	1	2	2	1	0
Four evaluators	1	0	2	0	1	3
All evaluators	1	0	2	3	5	2
>= Three evaluators	3	1	6	5	7	5

The degree of variation in the number of problems identified by single and multiple evaluators across the interfaces and applications (Figure 8.9) relates to the *evaluator effect*. The effect is not uncommon in the heuristic evaluation method, as noted by Hornbaek and Frojaer [2008] and Hertzum and Jacobsen [2003] and discussed in section 4.2.3.1.7. The variation suggests that a minimum of three evaluators (but preferably five) should be used for a heuristic evaluation. In this study, the upper limit of five evaluators for this minimum was used. Table 8.3 indicates that only 27 (38.02 percent) of the 71 problems would have been identified on average if only three evaluators were used.

The variation in the number of problems identified by the evaluators could also be related to their expertise in identifying usability and/or direct accessibility-related problems, as will be discussed in section 8.3.3.1.

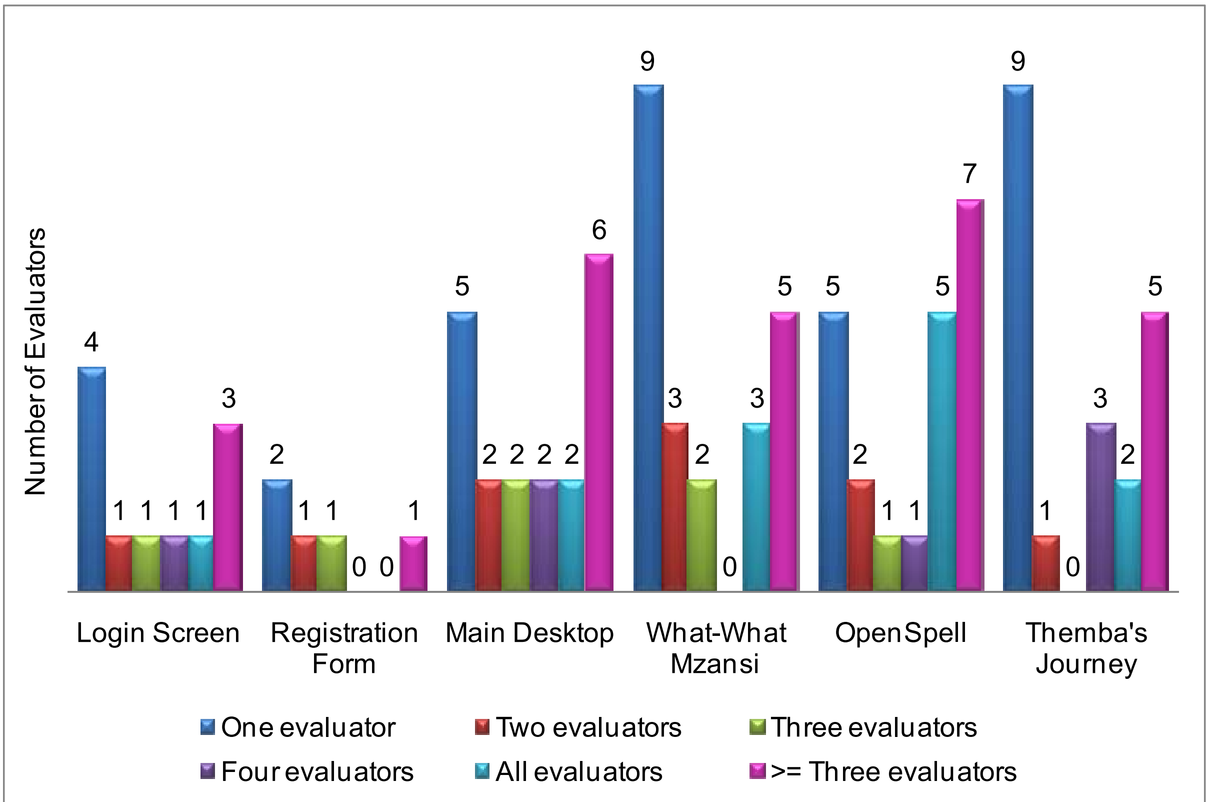


Figure 8.9: Variation in number of problems identifed per interface/applications

8.3.3 Problems Identified per Heuristic Category

8.3.3.1 Number of Problems Identified by Evaluators per Heuristic Category

The number of problems identified by each evaluator was analyzed according to the four heuristic categories. These numbers are shown in Table 8.5 and graphically illustrated in Figure 8.10. Some of the identified problems violated multiple heuristic categories and were classified based on the

Table 8.5: Number of problems identified by evaluators per heuristic category

	Heuristic category			
Evaluator	**General Usability Heuristics**	**Form Usability Heuristics**	**Direct Accessibility Heuristics**	**Game Usability Heuristics**
Evaluator 1	22	1	5	9
Evaluator 2	23	1	7	12
Evaluator 3	21	2	8	11
Evaluator 4	22	0	5	15
Evaluator 5	21	1	8	18

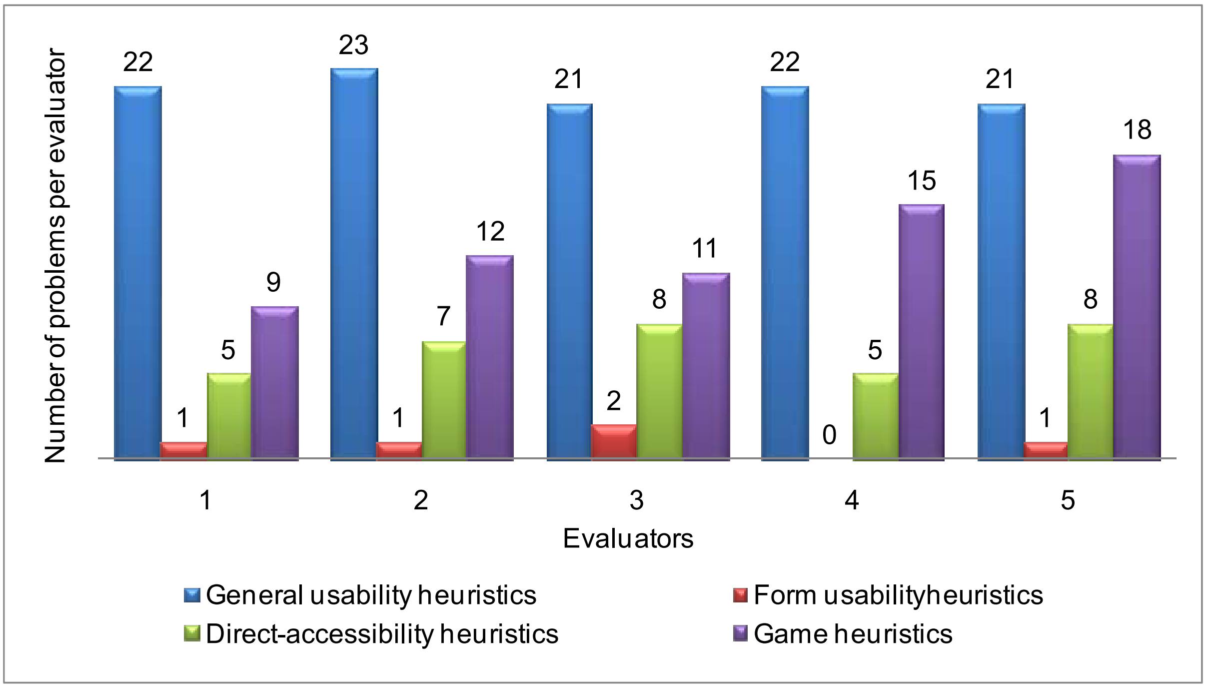


Figure 8.10: Number of problems identified by evaluators per heuristic category

heuristic category violated. There is a tendency for evaluators to detect more problems in one heuristic category than in another.

The five evaluators identified a similar proportion of general usability problems (between 21 and 23), although not necessarily the same set of problems.

It was not surprising that Evaluator 1 and 4 identified the lowest number of direct accessibility-related problems, as their main experience is in usability evaluation unlike Evaluator 3 and 5 have who have been involved in accessibility evaluations in the past. Evaluator 2 identified a relatively high number of direct accessibility-related problems, perhaps because this evaluator manages a usability testing laboratory where participants sometimes include those with disabilities, creating an increased awareness of accessibility issues. The results show the value of including evaluators with accessibility expertise in the evaluation of the DD.

Table 8.5 and Figure 8.10 also showed that Evaluator 4 and 5 identified fifteen and eighteen problems that violated game usability heuristics respectively. Evaluator 4 has been involved in various forms of e-learning over an extended period of time, which might explain the significant number of game related problems identified. However, no apparent reason could be given for the high number of game-related problems identified by Evaluator 5.

8.3.3.2 Total Problems per Heuristic Category

Analysis showed that 43 (60.56 percent) of the 71 aggregated problems identified violated general usability heuristics, while only three (4.23 percent) related to the form usability heuristics (see Figure 8.11). Fifteen (21.12 percent) problems violated direct accessibility heuristics while 26 (36.62 percent) were related to game usability heuristics.

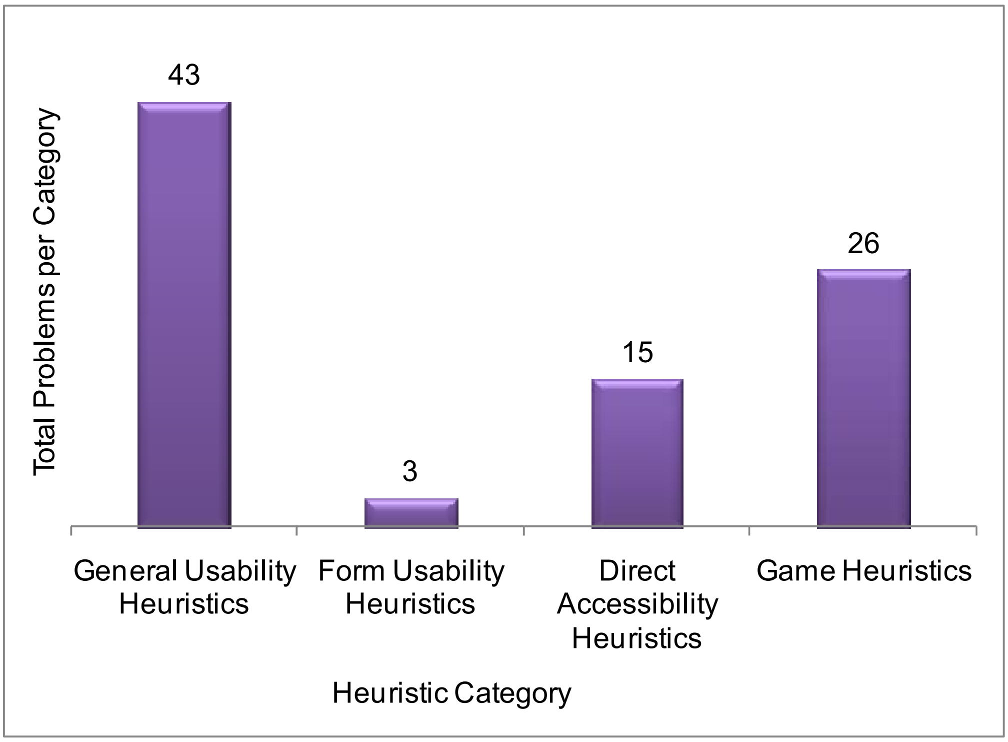


Figure 8.11: Total problems identified for each heuristic category

Because identified problems that violated more than one heuristic category were analyzed separately according to the heuristic categories that were violated, the numbers provided in Figure 8.11 does not yield the total 71, but rather represent the total numbers of problems identified per heuristic category.

As shown in Figure 8.11, more than half of the aggregated list of identified problems (60.56 percent) related to general usability heuristics. The high number of problems that violated general usability heuristics can impact on the achievement of one of the goals of the DD project, which is the promotion of computer literacy through unassisted learning [Cambridge, 2008]. Applications aimed at promoting self-learning should adhere to basic usability principles.

Figure 8.11 also shows that only three of the identified problems violated the form usability heuristics. This relatively small number was not surprising since the DD utilizes a simple form to collect data about new user accounts and I did not anticipate identifying significant number of problems in the form.

The relatively high number of problems violating direct accessibility heuristics was also expected, since the DD was not designed with accessibility in mind. However, compliance with the direct accessibility heuristics will go a long way in improving general usability for users with no apparent disability, since the environment of use of the system (for example, noisy surroundings, glaring sunshine) can impose some limitations on users.

With regard to problems violating the game usability heuristics, many of these were also general usability problems. Correcting the general usability problems will also address many of the game-related problems.

8.3.4 Locations of Usability/Accessibility Problems

An analysis of the location of the aggregated list of usability and direct accessibility-related problems (Table 8.6) revealed that the registration form alone recorded the least number of problems and *What-What Mzansi* the highest. However, a total of 21 problems were found within the login screen and the main desktop. The combined number of problems within these two interfaces is more than that found in any other interface or application evaluated. These two interfaces constitute the first contacts that all users have with the DD, whether they are registered users or guests, the latter of which represents the majority of users. The problems related to these interfaces are therefore significant in that they could affect the majority of users.

Figure 8.12 graphically illustrates the spread of the problems identified per interface/application relative to the aggregated list of problems.

Table 8.6: Number of problems per interface/application

Interface/Application	No of Problems
Login screen	8
Registration form	4
Main Desktop	13
What-What Mzansi	17
OpenSpell	14
Themba's Journey	15

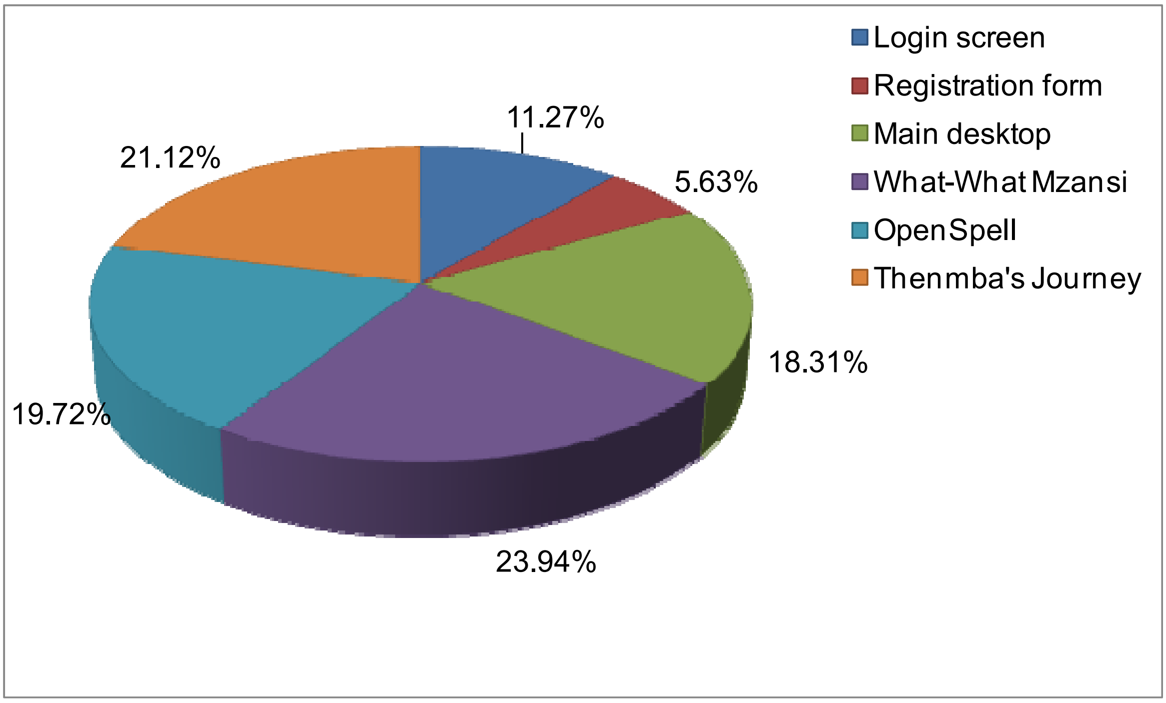


Figure 8.12: Problem percentage per interface/application

8.3.5 The Nature of Usability and Direct Accessibility Problems Found in the Various Application

The kinds of usability and direct accessibility problems found by expert evaluators have the potential to affect users' interaction with the DD. Many of these problems can classify as high-severity problems; for example, absence of feedback to users following the provision of an incorrect username and/or password. This section highlights the major usability and direct accessibility problems in the interfaces and applications evaluated.

8.3.5.1 Problems Related to the Login Screen

The login screen is the first DD interface experienced by users of the DD. One of the major usability problems relating to the login screen is the absence of feedback to users following the provision of an incorrect username and/or password, with the same screen returned over and over when the user cannot overcome the problem. The provision of appropriate and timely feedback is an essential usability principle that enables a user to determine the effect of an action that has been initiated, allowing him/her to determine what follow-on action to take (section 4.3). An experienced user may realize after a second login attempt that the information provided is probably incorrect. Given that the DD is targeted at users with little or no computer literacy, it may be difficult for such novice users to associate the return of the same screen with an incorrect username and/or password. Such user may well think that the system is not functioning properly.

Another problem with this interface is the absence of a mechanism for users to retrieve a forgotten password. The accepted standard for systems that require user authentication, prior to accessing the system's functionalities, is to enable users to retrieve previously stored password by answering simple questions that have been specified by the user at registration.

Although users benefit from accessing Internet services on the DD as registered users, the login screen did not provide any information to inform users' decision about accessing the DD as guest or registered user.

The main direct accessibility problem on the login screen in its current form relates to the font size of the login instructions provided in the IsiXhosa, Afrikaans, Sotho, and Venda languages. The small size of the instructions may constitute accessibility barrier to people with low visual abilities.

8.3.5.2 Problems Related to the Registration Form

The main usability problem affecting the registration form involves the close proximity of the <Register User> and <Cancel> buttons, and the absence of an alert message urging users to confirm a click on the <Cancel> button. Because of the close location of the buttons to one another, users (expert or novice) could easily click on the <Cancel> button while intending to click on the other. The DD is intended for users to be able to learn the use of the system with little or no assistance from an external party; thus, it should support users' exploration and their reversal of any unintended action.

8.3.5.3 Problems Related to the Desktop

The main desktop exhibits several major usability problems that may negatively affect users' interaction with the DD. One of the major problems involves the use of unintuitive icons to represent system functionalities. As shown in Figure 8.3, a green slider bar provides the volume control and a red right pointing arrow represents the logout button. Interpreting the functionalities of these two icons based on their appearance will be difficult for both experienced and novice users.

One of the principles of usability is the use of meaningful names for interface elements. Another problem with the desktop relates to the labels of icons and folders on the desktop. For example, the educational games evaluated in the study are accessed only by clicking on the folder labelled 'new_content', which does not suggest to users which applications are stored in the folder.

8.3.5.4 Problems Related to What-What Mzansi

The educational game application, *What-What Mzansi*, displays a number of general usability and game usability problems. The main general usability problem involves the icon that is used to provide context-specific instructions to users. The use of question mark '?' to represent an icon that provides important information about the application is problematic, considering that the DD is aimed at novices. Designers cannot reasonably assume that users with little or no computer literacy will be able to associate the icon with the function that it provides.

With regard to game usability, *What-What Mzansi* asks questions without users explicitly choosing to begin the exercise. Allowing users to indicate when they are ready to start the exercise will give them time to explore and familiarize themselves with the functionalities provided by the application. Another problem involves using sarcastic terms to provide performance feedback to users; for instance, feedback such as "don't make me laugh" will potentially demotivate a sensitive user.

8.3.5.5 Problems Related to OpenSpell

The major problems in *OpenSpell* relate to general usability and direct accessibility problems. Application menus are given names that are not meaningful, and the symbols used are unintuitive. Naming the menu that provide the words to be learnt <say> and that for spelling exercises <spell> will not allow users with little or no computer literacy to associate the labels with the functionalities that they provide. Also, the unintuitive use of stars (*) to represent the levels of difficulty will make it difficult for users to adjust a lesson's difficulty based on their ability and performance.

Another problem involves the unnecessary restriction to use the onscreen keyboard and the touchpad to provide input while completing spelling exercises. Apart from being a nuisance to the able user, users with limited use of their hands will find it difficult to move the pointer around on the screen to select letters of the alphabet.

8.3.5.6 Problems Related to Themba's Journey

Themba's Journey demonstrates the worst of the design decisions that were made by the developers of DD applications. Although the application supports the IsiXhosa and English languages, a non-IsiXhosa user is expected to hover the pointer over speech bubbles before the application can display the English equivalent of the story. One cannot imagine why the developers assumed that any user (experienced or novice) will know that they must do this. Not even the information provided in the 'Help' menu specifies this operation.

Another problem with the application reveals inadequate testing of the application before deployment. Although the application provides an <Exit> button on every screen and the main interface (except for the 'Help' menu screen), none of these buttons are functional.

8.3.5.7 Subset of Problems Identified by Evaluators

As discussed in section 8.3.2, the expert evaluators identified an aggregate of 71 problems in the DD. In addition to identifying problems, the evaluators also provided recommendations for correcting the problems. Appendix F presents the complete list of the identified problems, the heuristics that were violated, the evaluator(s) that identified the specific problems, together with the recommendations for improvement. Table 8.7 shows a subset of the list: including, categories of the heuristics (see section 7.3) that were violated by the specific problem; the number of the heuristic relative to its category; the description of the identified problems; and, the number of evaluators that detected the problem.

Table 8.7: Extracts from problems identified by evaluators

Heuristic category	Heuristics violated	Problem description	Identified by (Evaluator)
Login screen			
1	4.1	There is no feedback whatsoever when an incorrect username and/or password is provided.	E1, E2, E3, E5
3	1.1	The font sizes of the instructions on how to login/create user account in four other languages are small.	All
New account registration form			
2	3.2	There is no indication of which fields must be filled and which ones are optional.	E2
3	4.1	User cannot use the <Tab> key on the keyboard to select female for the gender field.	E3
Main desktop			
1	2.2	The functionality of the ⇒ button, used to exit the DD is not clear from its look. This button is also hidden from users' view.	All
1	5.2	The locations of the following icons on the taskbar are too close to one another: ⇒ button <System> <Volume control>, and Volume control slider. Users can easily click on the ⇒ button while trying to use the volume control slider, thereby closing the system unintentionally.	E5
3	3.2	The level of contrast between the dark blue background and the grey foreground used to label icons is low. The contrast between the word 'Digital Doorway' and the dark blue background is poor.	E3, E 5
What-What Mzansi			
1	3.2	At the start of the application, some of the control buttons	All

		and the character that reads out instructions and questions are hidden from user's view. A full screen mode is activated by clicking arbitrarily around the taskbar.	
3	4.1	The <PgDn> and <PgUp> buttons on the keyboard cannot be used to navigate the content of the window within the 'About' menu option.	E5
4	4.2	The performance feedback "don't make me laugh" after a poor performance is cheeky and not encouraging. Some users might find it offensive.	All
4	6.2	The volume control button is not visible when game is in full screen mode.	E1
OpenSpell			
1	2.2	The use of the labels 'Say', 'Guess', and 'Spell' are not descriptive of their functionalities.	All
1	2.2	The ***** symbols used to represent the level of difficulty are not intuitive.	All
3	1.5	The quality of the speech output is poor and not easily discernible even when the volume is at maximum level.	All
1 3 4	10.1, 10.2 4.1 2.3	When the 'Spell' option is selected, the user cannot use the keyboard to provide input but must use the onscreen keyboard.	E1,E 2, E3, E4
Themba's Journey			
1 3 4	1.1 1.2 2.1	To access an English version, the user must hover the pointer over the speech bubbles. This is problematic for users with limited use of their hands. The instructions provided under 'Help' do not specify this. No user will know s/he must do this until told.	All
1	8.1	*Themba's journey* provides an <Exit> button to close the application, while the same functionality is provided by the <X> button in *What-What Mzansi* and *OpenSpell*.	E5

As shown in Table 8.6, some of the identified problems cut across multiple heuristic categories. Consider the fourth problem described under *OpenSpell* "When the Spell option is selected, the user cannot use the keyboard to provide input but must use the onscreen keyboard". The problem violates the following general, direct-accessibility and game usability heuristics:

- The Digital Doorway should not impose unnecessary constraints on the user input method.
- Where user input can be provided via the keyboard and onscreen keys, the user should be allowed to provide input through either method.
- Allow keyboard navigation for operations/tasks that do not essentially require the use of the mouse.
- Clearly specify constraints and restrictions governing the game.

8.4 CONCLUSION

This chapter is directly linked to the conclusion phase of the outer cycle of the design research and provides part of the answer to my third research sub-question (What is the result of applying suitable method(s) to the DD?). The chapter started by describing the interfaces and applications evaluated and detailing the formal heuristic evaluation process. This was followed overviewing and discussing the results obtained from the heuristic evaluation.

The heuristic evaluation yielded a large number of usability and direct accessibility-related problems. Initial analysis involved the consolidation of similar problems. This process produced an aggregated list of 71 usability and direct accessibility-related problems.

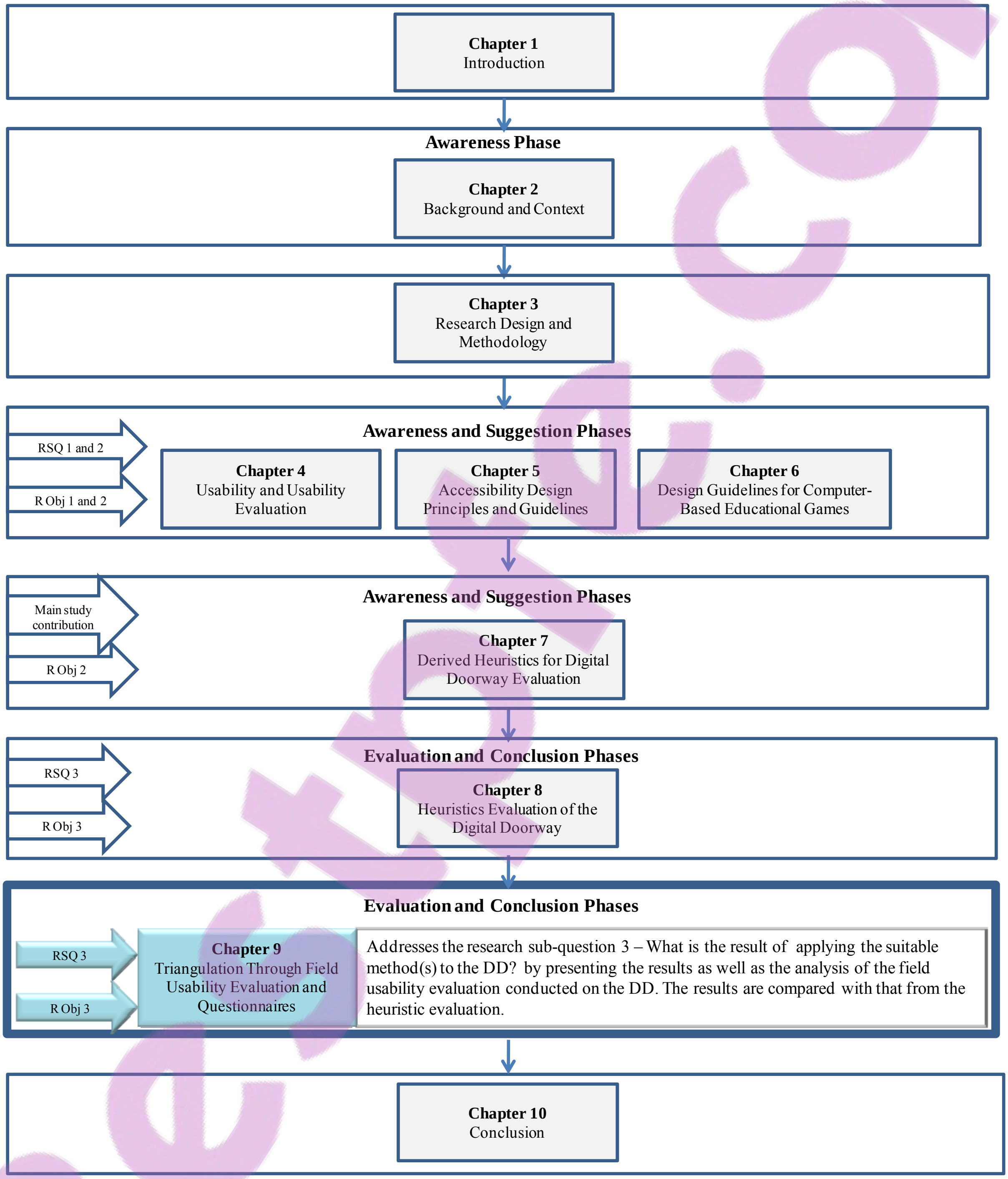
Analysis of the aggregated list of problems showed a high level of variation in the number of evaluators that detected specific problems, with 48 percent of the aggregate problems recognized by single evaluators. This indicated that the *evaluator effect* was indeed evident in the study.

The analysis of the result also showed that 60.56 percent of the aggregated list of problems violated general usability heuristics. This demonstrates the need for the formulation of basic usability standards that should be adhered to by the developers of DD applications.

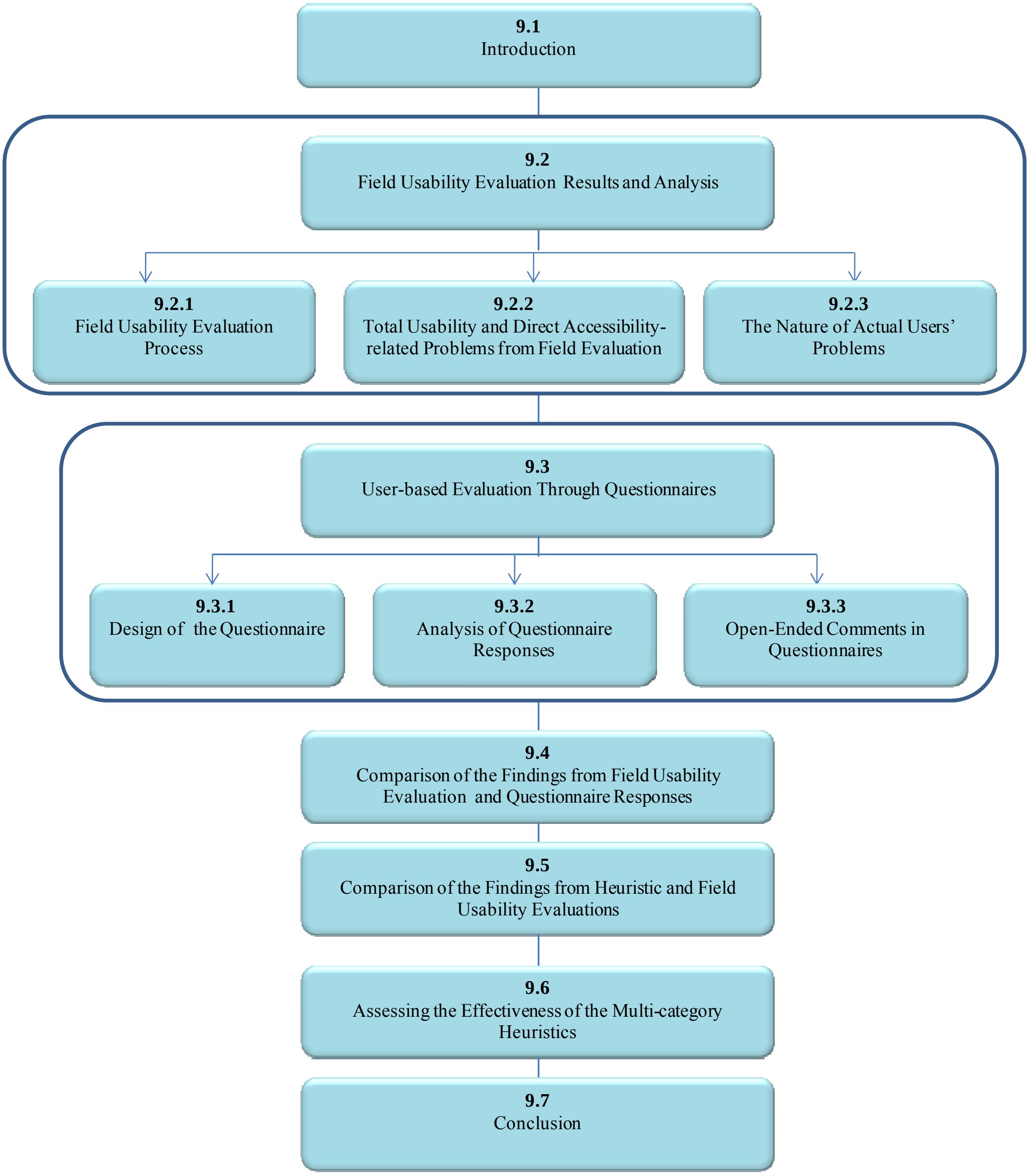
In the next chapter, I will discuss the results of the field usability evaluation and user evaluation through questionnaires, conducted as methods for data triangulation.

CHAPTER 9: TRIANGULATION THROUGH FIELD USABILITY EVALUATION AND QUESTIONNAIRES

The stage of Chapter 9 in the dissertation



Map of Chapter 9



9.1 INTRODUCTION

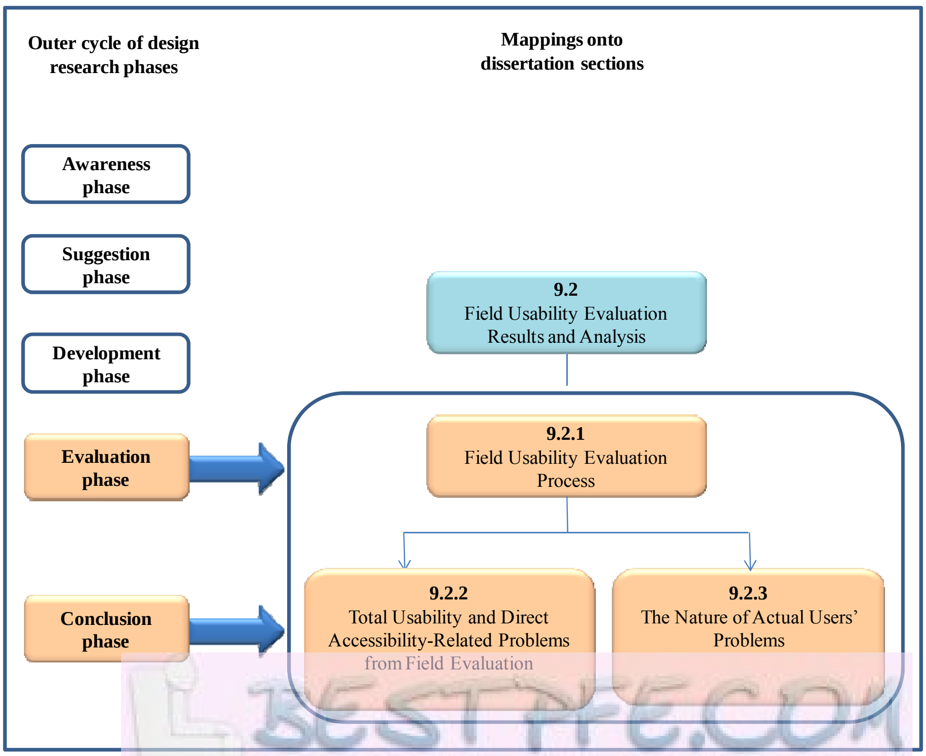
This chapter also maps onto the evaluation and conclusion phases of the outer cycle of the design research. Chapter 9 provides the remainder of the answer to the research sub-question, partially answered in Chapter 8: What is the result of applying suitable method(s) to the DD? The purpose of Chapter 9 is twofold: it serves as a triangulation exercise and an assessment of the complimentary role of various usability evaluation methods. Two triangulation activities were conducted; a field evaluation through direct observation to triangulate the heuristic evaluation process and questionnaires to triangulate data obtained from the field evaluation. The direct observation and user evaluation through questionnaires to evaluate the DD from users' perspectives were conducted at a local school.

In section 9.2, I provide the results obtained from the field evaluation by discussing the total number of actual problems encountered by users and the description of the problems. In section 9.3, I present analysis of users' responses to the questionnaire. Section 9.4 compares the results of the field evaluation with participants' responses to the questionnaire while section 9.5 provides the comparison of the results of the field evaluation with the problems identified by expert evaluators.

In section 9.6, I assess the effectiveness of the multi-category heuristics using the criteria suggested by Sim et al. [2009]. In addition, a variation of the comparison method by Ling and Salvendy [2005], and the nature of usability and direct accessibility problems identified by expert evaluators were also used to assess the heuristics. Section 9.7 concludes the chapter.

9.2 FIELD USABILITY EVALUATION RESULTS AND ANALYSIS

Mapping of section 9.2 to the design research phases



In general designers make a number of assumptions about the potential users of a proposed application or system [Gardner-Bonneau, 2010]; such as the potential user's educational background and the level of expertise. Evaluation with real users enables evaluators to assess the extent to which those assumptions are valid. This section discusses the procedure followed in the field usability evaluation to evaluate the usability of selected interfaces and applications on the DD from users' perspectives.

9.2.1 Field Usability Evaluation Process

A field usability evaluation was conducted at a local secondary school with the primary objective of triangulating the data obtained from the heuristic evaluation process. The decision to evaluate the DD at a school among other potential centres in which the DD is installed such as community centres and police stations was influenced by two factors:

1. The three applications evaluated were educational games; hence it makes sense to evaluate the usability of the applications among school children.
2. A study on the usage patterns of the DD at a number of representative centres around South Africa recorded secondary schools as having the most successful usage [Gush and De Villiers, 2010].

The direct field observation evaluation method was used to assess the DD from users' perspective in a natural environment of use. At the school, the DD is installed in an open area along a corridor to provide unrestricted access to users. A tarpaulin hanging from a wall provides some shading from the reflection of the sun. Children from surrounding homes also have access to the DD as soon as the school closes until 18:00 in the evenings.

To avoid disrupting learning activities as much as possible, the evaluation sessions were scheduled in the afternoons, after school hours, over a two-week period. The open area used for the evaluation meant that evaluation sessions commenced about 45 minutes after the school had closed to allow majority of the learners to disperse and minimize ambient noise in the session.

A pilot study preceded the field evaluation with three participants and each participant completed tasks involving the use of one of the educational games, *What-What Mzansi, OpenSpell* or *Themba's Journey*. Each evaluation session lasted between thirty and 45 minutes. The sessions were recorded on video cameras after assuring participants of their anonymity, and important events were noted in writing as they occurred. Each participant was also given a questionnaire (see section 9.3) to complete after the sessions. The questionnaires were completed and returned by the pilot participants the same day after the evaluation sessions.

The main field evaluation was conducted with nine learners participating. The evaluation process differed slightly from the conventional field study. Rather than simply observing participants while using the DD, pre-defined tasks were given to six of the nine participants using one of the applications, *What-What Mzansi*, *OpenSpell* or *Themba's Journey*. The use of pre-defined tasks enabled me to focus the evaluation on the selected interfaces and applications. To get a sense of the typical application usage by users at this particular center, three participants were free to select any application they wished to interact with within an allocated time of 45 minutes.

A cooperative evaluation style was also used in which I interacted with participants during each session. Rather than leave participants to 'wander' unassisted while completing tasks, they were encouraged to ask questions and seek assistance whenever they got stuck with any activity. This approach is justifiable since the DD is not a transaction processing system, where the speed at which tasks are completed is an important measure of usability. The approach also enabled participants to learn more about the functionalities of the DD and the specific application they used for the evaluation. The sessions were video-recorded and notes of significant actions/events were taken. The profiles of the main field study participants are provided in Table 9.1.

Participants with disabilities were not included in the field evaluation because the DD does not support the use of assistive devices, such as a screen reader for blind user.

As shown in Table 9.1, there were six male and three female participants. The age of the participants ranged between thirteen and eighteen years. Three participants were in Grade eight, one in Grade ten and two were in Grade eleven. The grades of three participants are unknown as their questionnaires were not returned (see section 9.3.2). This user group was selected as study participants since they represent the most prolific user group of DD applications [Gush and De Villiers, 2010].

Three participants used the application *What-What Mzansi*, two used *OpenSpell* and a further three used *Themba's Journey*. Other applications used by the participants who freely explored the DD were: *KTuberling* (a construction game), *Penguin* (a racing game) and *Four-in-a-row* (a logic game).

Observational data was documented in notes as well as video recordings of the sessions. This data was analyzed by first categorizing it according to the application used by participants, that is, *What-What Mzansi*, *OpenSpell* and *Themba's Journey*. Similar usability and/or accessibility problems were then grouped together before using descriptive texts to notate them.

During the evaluation sessions, the researcher observed that participants encountered various usability

Table 9.1: Profile of field study participants

Participants	Age	Gender	Grade	Application(s) used
1	17	F	11	OpenSpell
2	18	F	11	Themba's Journey
3	15	M	8	What-What Mzansi
4	13	M	8	OpenSpell
5	13	M	8	(Free Exploration) KTuberling and Penguin games
6	15	F	10	(Free Exploration) Themba's Journey
7	16	M	-	What-What Mzansi
8	14	M	-	Themba's Journey
9	15	M	-	(Free Exploration) What-What Mzansi and Four-in-a-row game

and direct accessibility-related problems. These ranged from the inability of participants to locate the required application, to lost data due to the lack of error tolerance by the system. Some of the problems affected the completion of participants' tasks while others constituted sources of minor irritations to them. Although the primary objective of the study was to evaluate the usability and the level of direct-accessibility support in applications on the DD, two hardware-related usability problems concerning the keyboard were also uncovered during the field evaluation. Sections 9.2.2 and 9.2.3 discuss the findings of the evaluation sessions in more detail.

9.2.2 Total Usability and Direct Accessibility-Related Problems from Field Evaluation

During the evaluation sessions I observed that participants encountered a total number of 39 actual usability and direct accessibility-related problems. Thirty seven of these were software problems affecting task execution by participants and two were hardware related problems. As illustrated in Figure 9.1, four problems involved *What-What Mzansi* while six affected the login screen, the main desktop and *Themba's Journey* respectively. Seven problems involved *OpenSpell* and eight related to the registration form.

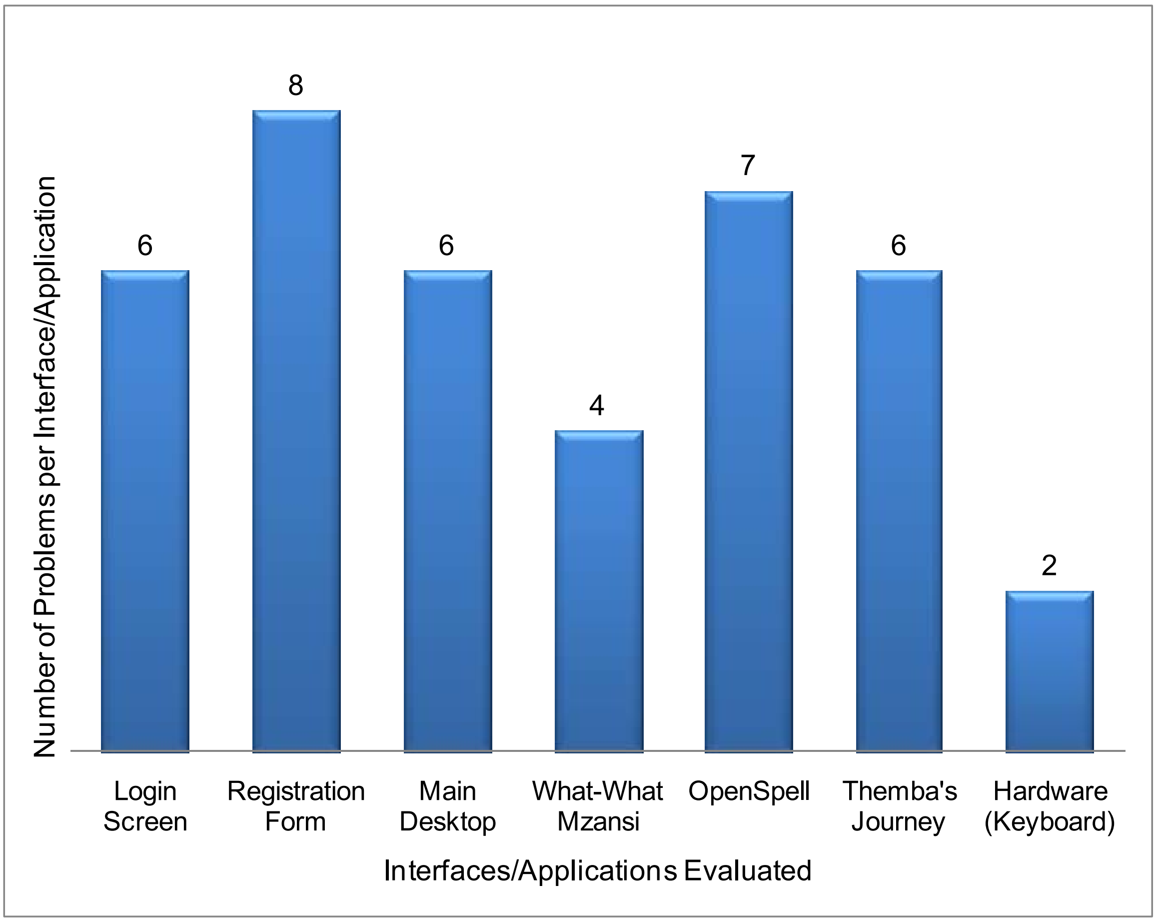


Figure 9.1 : Number of actual user problem per interface/application

Problems affecting the login screen, the registration form and the main desktop constitute 54 percent of the total software usability problems. At least two of these interfaces (the login screen and the main desktop) represent the first areas of contact between the user and the DD, and as section 9.2.3 describes, the problems can negatively impact on the ability of users to successfully interact with and explore the DD. This is a concern, since the DD aims to promote computer literacy through unassisted learning [Gush et al., 2010]. A system aiming for this kind of goal should provide simple and intuitive interfaces and allow system exploration.

9.2.3 The Nature of Actual Users' Problems

Nine secondary school students participated in the field usability evaluation and the software evaluated were educational applications, suiting this target group. Six participants undertook pre-defined tasks to focus the evaluation on the specific interfaces and applications and three were free to select any application they wished. The various tasks are described in each of the relevant sub-sections below. As part of the pre-defined task, the six participants were required to create new user accounts using the registration form, unless they had valid accounts. None of the six participants had previously registered as a DD user. The following sub-sections describe the problems experienced by participants within specific interfaces and applications.

9.2.3.1 Problems Relating to the Login Screen

Content in the DD can only be accessed after a successful login, either as a guest or registered user (section 8.2.1 and Figure 8.1). Eight participants created new user accounts, including all six participants performing pre-defined tasks and two who explored the DD freely. This process was initiated by typing 'new' in the username field on the login screen.

Only two of the eight participants creating new user accounts were able to log in with their username and password on first attempts and three participants successfully logged in at the second attempt. Three other participants had to log in as guest users due to the incorrect username and/or password they provided.

The following usability problems relating to the login screen were identified:

- After creating the user accounts the system returns to the login screen where users are required to enter their username and password to gain access to DD resources. However, participants did not know that they needed to enter the username and password chosen during the registration process. As the login screen did not provide any information to users who had just created new accounts or to those with existing accounts (Figure 8.1), some participants typed in 'new' or 'dd1' in the username field and others asked what they need to do next after spending a few minutes reading the information on the login screen. I had to intervene and inform the participants that they log in using the newly chosen username and password.
- Participants sometimes confused the terms 'surname' and 'username'. They typed their surname in the username field though this was not the chosen username at the time of registration.

- The DD provides no feedback after incorrect username and/or password but repeatedly returns the same screen again and again. Participants' faces made visible their confusion about what the problem was. After several failed attempts by three participants, I intervened and informed the participants that the username, the password, or both were not the ones selected during the registration process. I advised the participants to log in as guest users as the DD does not provide support for the retrieval of a forgotten username and/or password.
- A screen resolution dialogue box appeared and disappeared (after a few seconds) with the message "For best picture quality change the resolution to 1024X. 1: Exit 2: Delete." This sometimes frustrated participants as they did not know how to handle the information.
- On two occasions, when participants were about to place the insertion point in the username field, the following message appeared on a rollover: "Answer questions here and press Enter when done. For a menu press F10". While this message did not seem to bother the participants in this study, its relevance is questionable.
- The DD login screen does not provide any prompt or hint for users to press the 'Enter' key or an onscreen <Enter> button next to the username and password field. Some participants did know what they should do after entering their usernames in the username text box. Two participants specifically asked me for help and I had to tell another what to do after spending some time looking confused and unsure of the next required action.
- Some participants confused the 'Enter' key on the keyboard with the 'mouse click' keys above the touchpad because the keys were not labelled. However, after several presses on one without the desired effect they then pressed the other. This was identified as a hardware usability problem.

The login screen problems may have implications for the provision of Internet access through the GPRS technology, which is currently being explored in the DD project. This service requires users to pay for Internet access by loading prepaid mobile phone recharge vouchers onto the DD system and allows users to use any unutilized voucher balance at a later stage if they are logged in as registered users. Use of Internet service as a guest enables accessing the credit balance in a voucher loaded by a previous user; as users log in with the same 'name', such as 'dd1'. Thus, success in providing Internet access through the GPRS technology partly depends on users' ability to create a user account and remember the selected username and password. Currently, the DD does not support the retrieval of a forgotten username and/or password. Without the provision of this functionality, it is possible users may not use the GPRS service since they must pay for this service and they are not assured of being able to utilize any unused credit balance at a later stage.

9.2.3.2 Problems Relating to the Registration Form

Eight of the nine participants created new user accounts and variously encountered problems while completing the registration form:

- The insertion point is not located within the first data field at the start of the form. Some participants began typing their names immediately after the form was loaded, only to realize after a few key strokes that the input was not being accepted. This was because the form design requires users to position the insertion point in the first data field before typing.

- None of the form fields are indicated as being mandatory or optional. Participants typically left the home language and preferred language fields unfilled only to have error messages urging them to fill each field in turn.
- Some participants chose passwords with a length of fewer than six characters, despite a hint provided next to the password field: “6 to 10 characters”. This resulted in a contradictory error message: “Passwords must be between 6 and 14 characters”.
- While setting the password, a participant received the following error message: “The password contains illegal characters”. This participant could not comprehend the meaning of the error message and asked me for help.
- The form did not allow users to locate the fields with input errors quickly; for example, by highlighting the field. A participant erased his input in the password field accidentally, while trying to correct the name field entry following an error message. The insertion point remained in the password field after clicking on the <Register user> button. Without the participant realizing this, he pressed the ‘backspace’ key on the keyboard several times and erased the wrong field.
- Two participants accidentally clicked on the <Cancel> button. The <Cancel> button is located close to the <Register User> button (see Figure 8.2). This inadvertent user error resulted in the form being closed without any warning to the participants, thereby erasing all the information entered by the participants up to that point.
- Some participants input their name and surname in the “Full Name” field without inserting a space in between them. This common error will then bring up the following error message: “Your name seems to be incomplete”. Participants then spent some time trying to figure out what the problem was, sometimes without success until I provided them with hints on how to resolve the problem.
- Three participants were unable to delete the wrong input in form fields until I told them how to do it. This task could only be accomplished by pressing the ‘backspace’ key on the metal keyboard, which then deletes the input one character at a time with each press of the key. This was another hardware (keyboard) usability problem.

The registration form enables DD developers to automatically gather demographic and application usage pattern data on the DD. This data is typically transferred to a central server on a daily basis [Gush et al., 2010] but is only of value if accurate data can be collected. If users have difficulties in using the form, then there is little motivation to create own user accounts since content can easily be accessed through a guest log in.

The DD is aimed at people with little or no computer literacy and the DD utilizes a simple form, with only two main groups of user data required. However, the nature of problems that participants encountered showed a lack of adherence to basic usability principles in the registration form. Error messages are provided in technical terms and contradicted hints on the form and users are not informed that all data fields are mandatory. Participants who chose to leave some fields unfilled had to deal with error messages one after the other urging them to fill empty fields.

More crucially, the designers of the form did not anticipate or make provision for unintended user errors. The <Register User> and <Cancel> buttons are located close to each other on the registration form. The proximity of these buttons could easily lead to any user (novice or experienced) inadvertently clicking on one, while the intention was to click on the other. An accidental click on the <Cancel> button by two participants in the study led to the closure of the registration form, and loss of previously entered data without any warning message.

9.2.3.3 Problems Relating to the Desktop

Following a successful log in, DD content can be accessed by clicking on icons/folders on the desktop or by selecting from one of the two menu options "Programs" and "Resources". The volume of audio output can also be adjusted using the global volume control provided on the desktop (see section 8.2.3 and Figure 8.3). The specific problems encountered by participants relating to the main desktop were:

- Only two of the six participants undertaking pre-defined tasks, using one of the educational game applications, *What-What Mzansi*, *OpenSpell* or *Themba's Journey*, were able to locate the game applications independently. The other participants unsuccessfully searched for the applications within the "Game" submenu, which is located in the "Resources" menu. These applications are however located in the "new_content" folder on the desktop. After several failed attempts to locate the applications, I had to inform the participants where to find the applications. It is possible that the two participants who found the game applications on their own were part of the crowds of students that gathered around the DD during earlier evaluation sessions.
- Only three of the six participants undertaking pre-defined tasks located the volume control buttons on the desktop and the remainder required assistance after several failed attempts.
- After clicking on the required game application icon, the screen flickers and returns to the DD home page. Participants needed to click the application icon several times before the game application opened. This was frustrating to the participants.
- Four participants found the background colour of the main desktop to be too dark. On several occasions they had to shield their faces and the screen with their hands while using the DD to overcome the extent of reflection of the sun on the dark background. The dark background was significantly worse than that experienced in the closed-up laboratory used by expert evaluators. The reflection worsened the contrast issue.
- A participant accidentally clicked on the ⇒ button, used to exit the system, while trying to locate the volume control button and the system was shut down without any warning.
- Only three of the six participants undertaking pre-specified tasks logged out of the system without requiring assistance. One participant discovered the ⇒ button accidentally following an attempt to increase the volume of audio output and two participants specifically asked for help following

failed efforts to exit the system on their own. Of the three participants that explored the system freely, two knew the location of the ⇒ button while the other participant asked for help after unsuccessful attempts to log out on her own.

The nature of the problems experienced by participants attempting to access applications and use the resources provided on the main desktop again revealed the lack of adherence to basic usability design principles. The educational game applications, *What-What Mzansi*, *OpenSpell* and *Themba's Journey* were hidden inside a folder that was named with little indication of the type applications it contains. The icons used to access essential functions were not intuitive (for example, the 'volume control' and 'exit' icons).

One of the aims of the DD project is to promote computer literacy through unassisted learning [Gush et al., 2010]. Thus, the system should support exploration by users and have built-in mechanisms to guide against potentially 'disastrous' user actions. One of the participants in this study unintentionally clicked on 'exit' while searching for the volume control button. The system did not prompt the user to confirm that she indeed intends to log out of the system and shot down without any warning.

9.2.3.4 Problems Relating to What-What Mzansi

Three study participants used the quiz game, *What-What Mzansi*. Two of the participants used the application as part of the pre-defined task (Figure 9.2) and the other participant chose to use the application. The problems experienced by the participants while interacting with *What-What Mzansi* were:

- Neither of the two participants using this application as part of the pre-defined tasks was unable to access the game instructions, as required in the specified task. Both participants searched for the game instructions by clicking on the "about" menu option which provides information on the application developers and the DD project history and achievements.

Digital Doorway Evaluation: Task list

1. Read the screen instruction on how to register as a new Digital Doorway user if you are not a registered user.
2. Complete the registration form if you are not a registered user, otherwise proceed to step 3.
3. Start the Digital Doorway by providing the requested information.
4. Search for the quiz game '*What-what Mzansi*'.
5. Remember to provide verbal feedback all the time.
6. Search for and read the instruction on how to play the game.
7. Proceed to play the quiz game.
8. Choose how challenging (difficult) you want the game to be.
9. Change the volume to suit your need.
10. Remember to provide verbal feedback all the time.
11. Close the Digital Doorway when you are done.

Figure 9.2 : Task list for field usability evaluation (using the application 'What-What Mzansi')

- Participants never used the context-specific instructions provided through the <?> icon (see Figure 8.4).
- None of the two participants who used *What-What Mzansi* for the pre-defined task knew how to get the full screen view of the game. At the start of the application, some of the control buttons and the character that reads out the questions were hidden from users' view. A full screen mode of the application is activated by clicking arbitrarily around the taskbar. The participant who elected to use *What-What Mzansi* as part of the free exploration of applications on the DD was able to change to a full screen view without requiring any help.
- One of the terminals used for the evaluation sessions had unusually large icons (Figure 9.3) compared with the normal screen (Figure 8.5). This resulted in non-visibility of a number of control buttons. In this particular instance, the right pointing icon '>' used for forward progression, was hidden from users' view. This made it impossible for the participants to repeat the level which they had just completed as required after a poor performance.

What-What Mzansi is a general knowledge quiz application (see section 8.2.4). Although the interface of the application is simple, the choice of caption for the third menu option "about" was misleading. The participants expected to find the game instructions in this menu. In addition, the label did not follow industry standards and conventions. Typically, an 'about' sub-menu is used within a 'help' menu of applications to provide copyright and version information of a particular application.

None of the three participants who used the application clicked on the <?>, which is available during the question sessions and returns context-specific instructions. Question sessions begin immediately after the welcoming words by the program character without the user getting the opportunity to access the instructions. Possibly, the main priority of the participants is to listen to and try to answer the

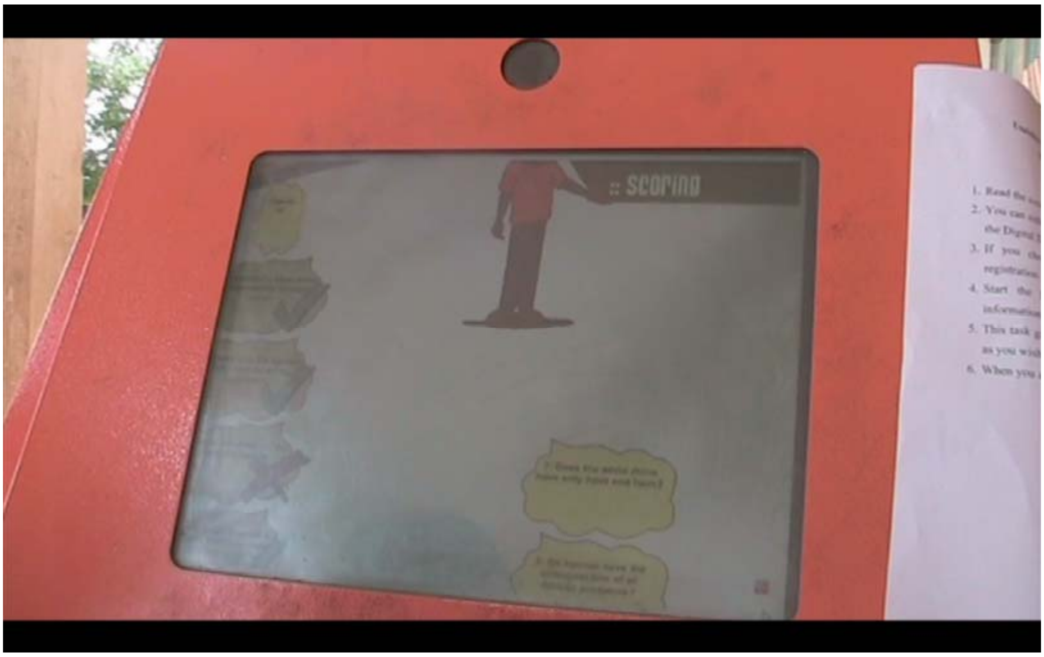


Figure 9.3: Screenshot from What-What Mzansi (with control buttons and part of character hidden)

questions as fast as possible since the timer is running. Accessing the context-specific instructions will 'eat away' at the time available. Another possibility could be that the participants simply do not notice it.

One example of a necessary context-specific instruction is a description of the mechanisms for providing answers to the questions. Questions can be answered most efficiently by pressing the left or right arrow keys on the metal keyboard, but none of the participants did so. Rather, participants used the touchpad to position the pointer over the correct 'pot' (see Figure 8.5) before pressing the 'mouse click' key above the touchpad.

9.2.3.5 Problems Relating to OpenSpell

Two participants used *OpenSpell*, an educational spelling application, for the second pre-defined task (Figure 9.4) in the field usability evaluation. The problems experienced by these participants were:

- Similar to the participants who used *What-What Mzansi*, the two participants that used *OpenSpell* were unable to access the game instructions. Both participants also searched for the instructions in "about" menu option, this menu also contained information on *OpenSpell* developers and DD project history and achievements.
- The two participants who used the application selected the "spell" option (see Figure 8.6) for task eight in the task list. However, the required functionality is provided within the "say" option.
- Only one of the two participants was able to associate the star (*) symbols with the level of difficulty. The other participant did not know how to set the difficulty level.
- The quality of the voice output was poor even when the volume was at the highest. Participants frequently had to keep their ears close to the screen. Although quality speech output is a must

Digital Doorway Evaluation: Task list

1. Read the screen instruction on how to register as a new Digital Doorway user if you are not a registered user.
2. Complete the registration form if you are not a registered user, otherwise proceed to step 3.
3. Start the Digital Doorway by providing the requested information.
4. Search for the spelling game '**OpenSpell**'.
5. Remember to provide verbal feedback all the time.
6. Search for and read the instruction on how to play the game.
7. Choose how difficult you want the game to be.
8. Learn how to spell a few words.
9. Change the volume to suit your need.
10. Do some spelling exercises.
11. Do a few guessing exercises.
12. Change the language to another one of your choice.
13. Close the Digital Doorway when you are done.

Figure 9.4 : Task list for field usability evaluation (using the application 'OpenSpell')

have for hard of hearing people, the poor speech quality affected the usability of *OpenSpell* for all participants because the DD was being used in an open environment at this particular center.

- For task ten in the task list, both participants first attempted to use the keyboard to provide their inputs, only to realize later that they could only use the onscreen keyboard.
- The terminal with the large icons, discussed in section 9.2.3.4, also affected task execution while a participant was using *OpenSpell*. Due to the large icons, the taskbar covered the three control buttons, <Repeat>, <Erase> and <Enter>, almost completely. On two occasions, the participant needed to erase incorrect inputs. Due to non-visibility of these buttons she clicked on the rightmost control button, which happened to be the <Enter> button. This was interpreted as an incorrect answer by the application. The participant was then offered a second opportunity to answer the question correctly. Figure 9.5 shows a screenshot from this terminal, which can be compared to the fully visible controls shown in Figure 8.6.
- The reflection from the sun affected the visibility of pictures displayed by the application. This was due do the inadequate provision of shading from the sun. The reflection can be seen in Figure 9.5.

The developers of *OpenSpell* replicated non-adherence to industry standard and convention in the use of "about" menu. As discussed above, an 'about' menu that gives the copyright and version information of a specific application is usually provided as a 'help' sub-menu.

Participants who used *OpenSpell* also searched for the game instruction in "about". This showed that

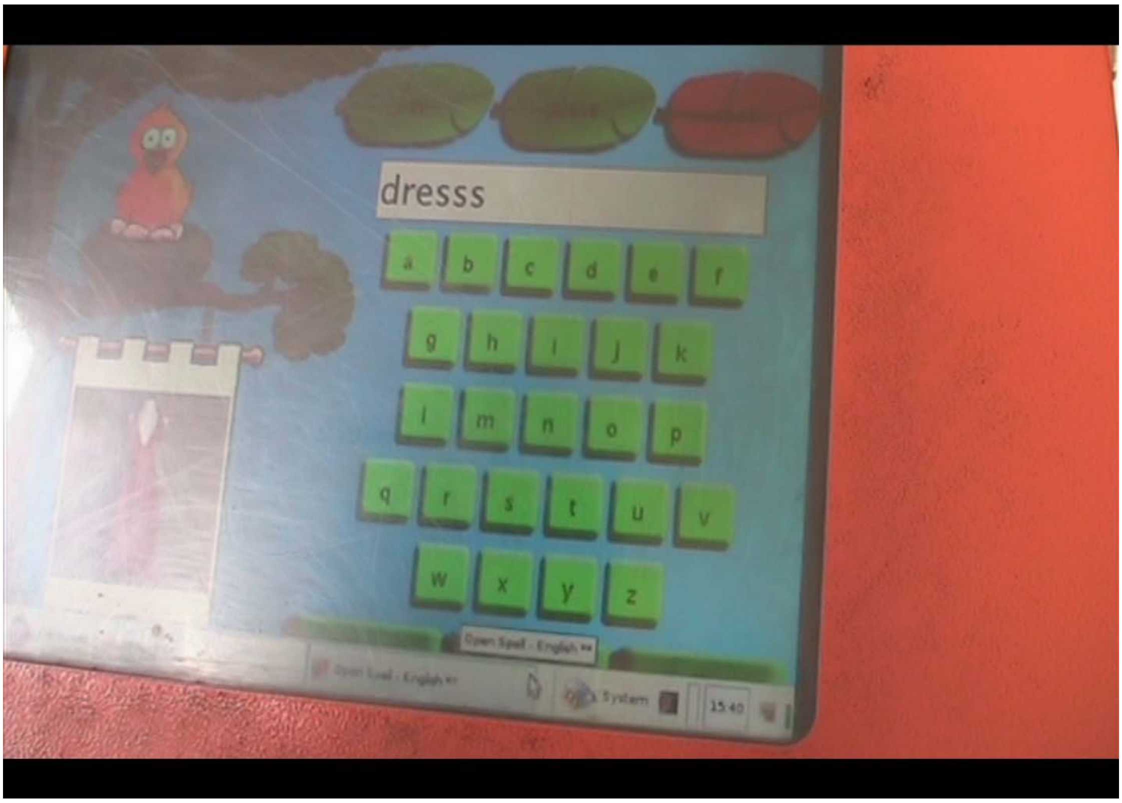


Figure 9.5 : Screenshot from OpenSpell (with control buttons obscured by taskbar)

this label can easily be interpreted as an application instruction, especially by users with limited computer literacy. Further, *What-What Mzansi* provided the instructions for the game through the <?> icon; by not providing any form of instructions or help for users, *OpenSpell* also lacked consistency with other applications on the DD.

Labels and symbols were also unintuitive. For example, the label “say”, provides the functionality of the lessons to be learnt; and stars (*) represent the level of difficulty of the application. As discussed above, only one of the two participants who used the application was able to associate the stars with the difficulty level.

OpenSpell restricts users to clicking on the onscreen keys on the application interface to enter input during spelling exercises and users cannot input using the keyboard. The first intuition of the participants who used the application was to press the keys on the keyboard.

The DD at this particular school was installed in an open area along one of the school’s corridor, a publicly accessible location. However, usability problems arise from the reflection of the sun on the system. I noticed the problem on my first visit to the school in the company of members of the DD team, and as a result a tarpaulin was hung from a wall to counter the reflection from the sun. The field evaluation has shown that the tarpaulin did not provide sufficient shading from the sun’s glare.

9.2.3.6 Problems Relating to Themba’s Journey

Three participants used the application *Themba’s Journey*. Two of these participants undertook the pre-defined the task (Figure 9.6) and the other participant elected to use the application. The problems experienced by participants were:

- The default language for *Themba’s Journey* is IsiXhosa and users must hover the pointer over speech bubbles to reveal English equivalents of the story (Figure 9.7). All the three participants who used the application had no idea how to get the English equivalent until I told them what to do.
- Too much physical effort was required by participants to move the pointer around the speech bubbles to read English versions. Their fingers became damp on several occasions due to constant movement over the touchpad to position the pointer.
- The application background was very dark. Participants shielded their faces and the screen with their hands. The dark background was made worse because the DD at the school was installed in an open space with excessive natural lighting and sun glare.
- The “Help” menu provides navigation instructions. Although the participants read the instructions at the start of the session, they tended to forget the functionality of some of these buttons. For example, when they were completing tasks nine and eleven (Figure 9.6), they forgot that the <Skip> button was designated by the double right pointing triangles ►► and that the <Review> button was designated by the double left pointing triangles ◄◄.

Digital Doorway Evaluation: Task list
1. Read the screen instruction on how to register as a new Digital Doorway user if you are not a registered user.
2. Complete the registration form if you are not a registered user, otherwise proceed to step 3.
3. Start the Digital Doorway by providing the requested information.
4. Search for the life-skills game '**Themba's Journey**'.
5. Remember to provide verbal feedback all the time.
6. Search for and read the instruction on how to play the game.
7. Proceed to play Themba's Journey.
8. After listening to Themba's Journey for a while, change to the English version of the story.
9. Change the presentation speed to view all the scenes available on the current screen.
10. Change the volume to suit your need.
11. Go back to the scenes of the previous screen.
12. Close the Digital Doorway when you are done.

Figure 9.6: Task list for field usability evaluation (using the application Themba's Journey)

- At the second crossroad, which had the two options "Walk"' and "Take taxi", the "Walk" option could not be executed. One participant had to select the "Take taxi" option against her wish.
- The main exit button was non-functional. Participants had to close the application with the browser exit button, that is, the <X> button (see Figure 8.7).

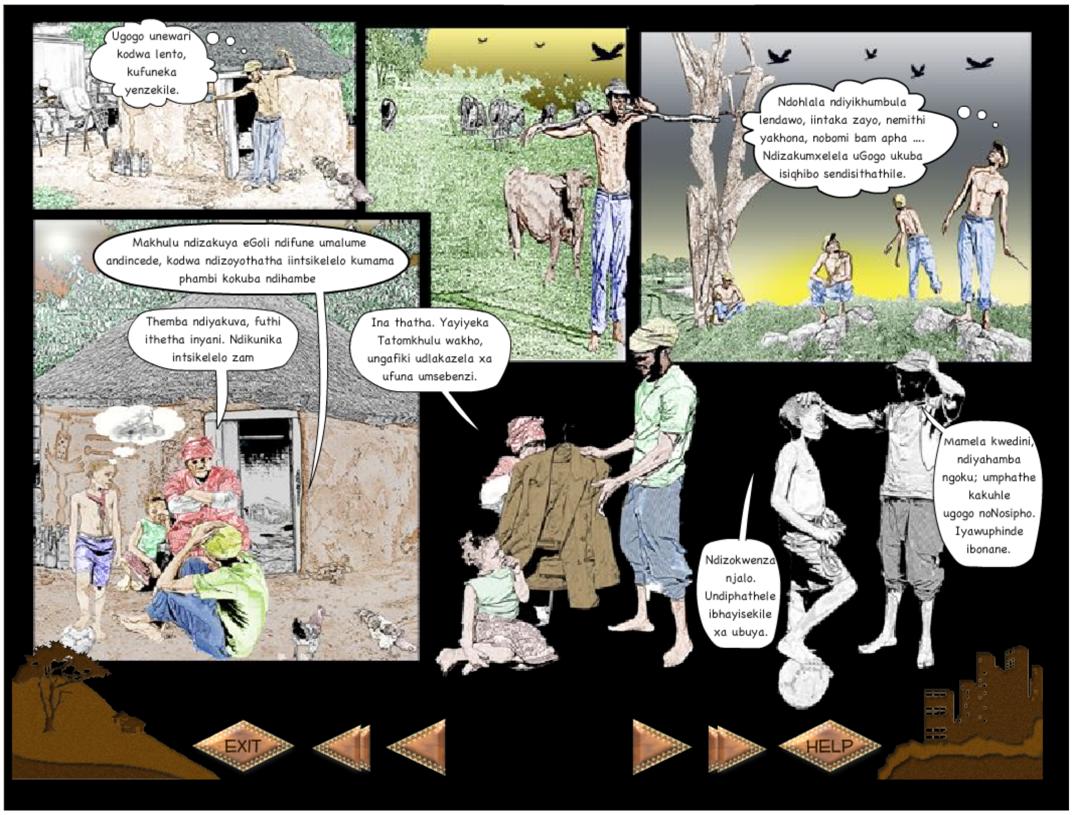


Figure 9.7: Screenshot from Themba's Journey

Themba's Journey is a life skills application which is embedded in scenarios that many young South Africans can relate to (see section 8.2.6). However, poor design decisions can limit the number of young people who can potentially benefit from the useful information contained in the application.

Although the application is available in both the IsiXhosa and English languages, the three participants were unable to access the English equivalent without assistance from the researcher. Also, frequent finger movements over the touch pad resulted in dampness, which the participants needed to dry on several occasions.

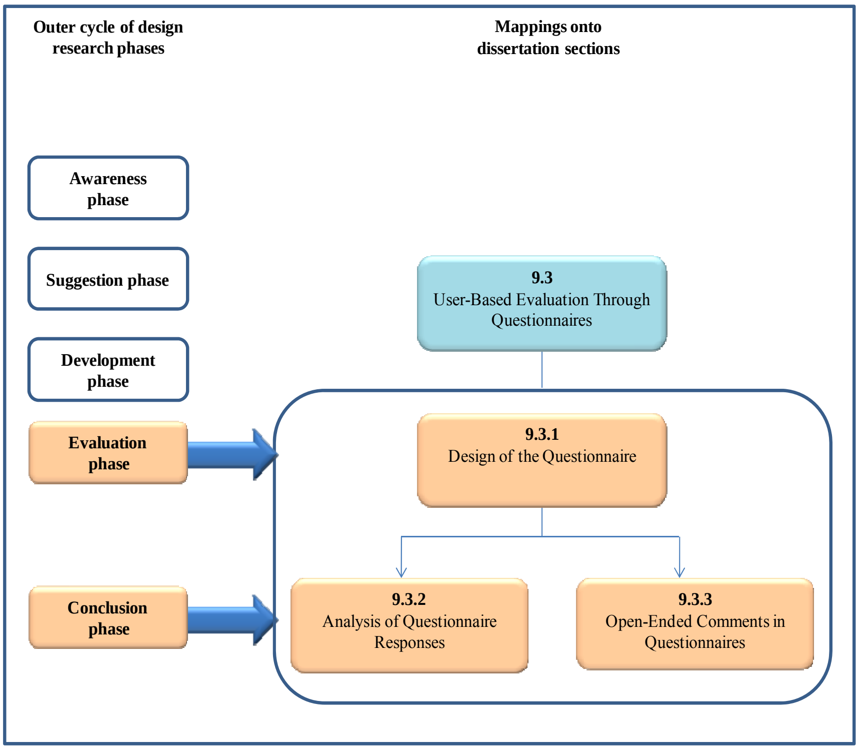
As discussed above, the lack of adequate shading meant that there was a constant glare of the sun over the DD screen. The dark background of *Themba's Journey* was aggravated by the reflection of the sun when using the application at the school.

The navigation instructions provided through the "Help" menu was read by the three participants after launching the application. None of them remembered the functionality of the <Skip> and <Review> buttons when they really needed to use them. However, the provision of "Help" button on every screen of the application proved to be beneficial as it allowed participants to easily look up the functions of these buttons.

Inadequate testing before the deployment of the application was also revealed by the inability of a participant to execute the "Walk" option at one of the crossroads in the application. In addition participants were unable to close the application using either the <Exit> button (see Figure 9.7) provided on each screen of the application, or from the main interface <Exit> button (Figure 8.7).

9.3 USER-BASED EVALUATION THROUGH QUESTIONNAIRES

Mapping of section 9.3 to the Design research



As a triangulation exercise to the user observations, a semi-structured questionnaire was used. The questions were based on a selection of the heuristics used by the experts. Section 9.3.1 discusses the design of the questionnaires, while sections 9.3.2 and 9.3.3 analyzes the responses from the completed questionnaires.

9.3.1 Design of the Questionnaire

As a user-based evaluation method, a selection of the heuristics generated for heuristic evaluation was converted to questions (in the form of positive statements) using simple terminology that can be comprehended by novices. Participants provided ratings of the level of their agreement or disagreement to these statements using a five point Likert scale: *Strongly agree*; *Agree*; *Neither agree nor disagree*; *Disagree;* and, *Strongly disagree* (see Table 9.2). In addition, sufficient space was provided after each statement where participants could write the specific problems they experienced during the evaluation which applied to that statement.

Each participant was required to complete the questionnaire after interacting with the DD. Because evaluation sessions were conducted in the afternoons, participants requested that they take the questionnaires home to complete and return the following day. This request was obliged because to avoid inconveniencing the learners as much as possible; moreover, participation in the study was voluntary. Of the nine participants, four returned the questionnaires the day after their evaluation sessions. One returned the questionnaire three days after the evaluation while another returned it several weeks after the evaluation. Three questionnaires were never returned.

Questionnaires that were returned by the participants were analyzed by summarizing the ratings of the participants to specific statements in the questionnaires. Open-ended comments provided by respondents were read, taking note of those that related to similar problems. The comments were recorded verbatim as written by the respondents.

Taking into account data obtained from the evaluation sessions, the ratings and comments of participants to the questionnaire were then compared with their behaviour during the evaluation sessions to corroborate and clarify the ratings and comments.

9.3.2 Analysis of Questionnaire Responses

Only six of the nine participants in the field evaluation returned the questionnaires. Table 9.2 provides the summary of the ratings of the six participants who returned the questionnaires. Columns 1 to 5 represent the ratings of the six participants that returned their questionnaires for each of the statements (numbered 1-23), while column 6 indicates the number of participants who provided no ratings for the specific statement. For example, the ratings of the participants to statement one “Instructions about how to use the Digital Doorway are clear to me” were: *Agree* (two participants); *neither agree nor disagree* (two participants); *disagree* (one participant); and, *strongly disagree* (one participant). To statement ten, “The Digital Doorway contains words used by computer people which I do not fully understand”, the ratings were: *strongly agree* (one participant); *neither agree nor disagree* (one participant); *disagree* (two Participants); *strongly disagree* (one Participant); with one participant providing no rating for the statement.

Table 9.2: Summary of questionnaire ratings

Strongly agree	**Agree**	**Neither agree nor disagree**	**Disagree**	**Strongly disagree**	**No response**
General ease of using the Digital Doorway					
1.	Instructions about how to use the Digital Doorway are clear to me.				
	2	2	1	1	
2.	Instructions and information are given in various ways like written words, spoken words, and through pictures.				
	4	1	1		
3.	I can change the volume of spoken words and instructions to suit me.				
2	3		1		
4.	Whenever I make a mistake, the Digital Doorway tells me exactly what the mistake is in a way that I can understand.				
2	1	1	2		
5.	The Digital Doorway clearly shows me what I need to do next so that I can complete what I am doing.				
1	3		2		
6.	When I choose an item on the Digital Doorway, the information presented to me is what I expect it to be based on the title of that item.				
3	3				
7.	I understand the meaning of icons (pictures) used in the Digital Doorway.				
1	4	1			
8.	I am able to determine the meaning and purpose of signs and symbols used in the Digital Doorway.				
	5		1		
9.	I am able to carry out similar activities in the same way in different parts of the Digital Doorway (For example, I can choose the language that I prefer in the same way).				
4	1	1			
10.	The Digital Doorway contains words used by computer people which I do not fully understand.				
1		1	2	1	1
11.	The Digital Doorway shows me how to correct my mistake.				
1	2		2		1
12.	The Digital Doorway gives response within a short time.				
1	2		2		1
13.	Spoken instructions and information are loud and clear enough for me to hear and understand.				
	2	1	1	1	1

Ease of using the registration form in the Digital Doorway					
14.	It is easy for me to determine which information I must give and which ones I may leave.				
2	1	1	1		1
15.	The space provided is sufficient for the information I need to write.				
1	3				2
16.	I am able to determine where I should write information in the form.				
2	2				2
17.	When I leave out information that is needed the Digital Doorway informs me about the missing information in a way that I can understand.				
1	2	1			2
18.	It is easy for me to make corrections to information that I have written earlier without me having to retype all over again.				
1	2			1	2
Usefulness of educational games in the Digital Doorway					
19.	The Digital Doorway does not make fun of me when my answer is wrong.				
2	1			1	2
20.	It is easy for me to choose in which language I want to play a game.				
2			1	1	2
21.	The Digital Doorway informs what I should to in order to play the games.				
2			1		3
22.	I am able to control how easy or difficult I want the game to be.				
1	1	1	1		2
23.	It is easy for me to determine what the computer games in the Digital Doorway will be teaching me.				
2	1	1			2

Of the six participants, three participants did not provide a rating to several of the statements, and when they did provide a rating, the answer always varied between *strongly agree*, *agree*, or *neither agree nor disagree*. These positive responses were contradictory to what was observed during the evaluation sessions. For example, to the statement “I am able to determine the meaning and purpose of signs and symbols used in the Digital Doorway”, two of these three participants ticked *agree*. However, one of them was unable to locate the volume control and exit buttons without assistance, while the other one could not set the level of difficulty in *OpenSpell*.

The remaining three participants selected a rating to most of the questions and provided additional feedback to some statements.

Based on the ratings of the participants, as reflected in Table 9.2, it appeared the registration form has little or no usability problems associated with it, with many of the respondents agreeing or strongly agreeing to the six statements relating to the form. Only one participant recorded a *strongly disagree* rating to one of the statements regarding the form.

Also the statements relating to the intuitiveness of icons and symbols (statements seven and eight in Table 9.2) were rated highly, with four participants agreeing to statement seven, and five agreeing to statement eight. Only one participant disagreed with statement eight.

9.3.3 Open-Ended Comments in Questionnaires

Three of the six participants that returned their questionnaires also provided open-ended comments relating to some of the problems they experienced during the evaluation sessions. The comments by the participants are reproduced below:

- "The projection of the voice is very low even though you up the volume it is still the same".
- "The screen is very dark for you to see the Themba game".
- "It is not easy for me to shut down the Digital Doorway".
- "At times the instructions are not clear, especially if you are a new user and the words that are used don't explain everything clearly".
- "I can't change the volume because it is not written where is the volume you just have to find it yourself which is not good".
- To the statement "The Digital Doorway gives response within a short time" two participants provided the following comments:
  - "Sometimes the system is very slow and we go away".
  - "It takes time which we do not have. Something must be done".
- To the statement "I am able to determine the meaning and purpose of signs and symbols used in the Digital Doorway" the following comment was provided by a participant:
  - "No, they have to be written, not putting symbols and expect us to find where is what I mean that's not good".
- To the statement "It is easy for me to choose in which language I want to play a game" the following comment was provided by a participant who used *Themba's Journey*:
  - "It is hard for you to choose the language because you won't see where you should choose the language".
- To the statement "The Digital Doorway shows me how to correct my mistake" two participants commented:
  - "Not at all times, sometimes you don't even know you've made a mistake".
  - "Not all the time I sometimes have to get my teacher to help me".

When compared with the ratings to the statements in the questionnaire, the open-ended comments provided by the participants showed the degree of frustration the participants experienced while interacting with the DD, although their overall ratings to the statements were positive. Without the opportunity to provide their views on the usability of the interfaces and applications they used, the ratings on its own would have given the impression that the participants were generally pleased with the usability of these interfaces and applications. As stated in section 4.2.3.5, the problem of respondents providing answers which they deem to be ‘acceptable’ is one of the limitations to the use of the questionnaire method for evaluation. This was confirmed in this study.

9.4 COMPARISON OF THE FINDINGS FROM FIELD USABILITY EVALUATION AND QUESTIONNAIRE RESPONSES

This section compares the results obtained from the field evaluation (section 9.2) with those from user evaluation with questionnaires (section 9.3).

Evaluation through the questionnaire method on its own relies on participants to complete the questionnaires and return them. As discussed in section 4.2.3.5, the problem of low return rate is not uncommon when using this method. Even when the questionnaires are returned, the information provided by respondents may not provide a true reflection of the usability of the application being evaluated. As discussed in section 9.3.2, three of the six participants who returned their questionnaires did not provide any rating to many of the statements in the questionnaire.

The results from the questionnaires also highlighted the kind of discrepancies that may occur between participants’ responses and their actual behaviour during evaluation sessions. As stated in section 9.3.2, the responses to the questionnaire by two of the participants contradicted what I observed during the evaluation sessions.

In the field evaluation method, data on the evaluation sessions were available immediately because it was captured live, using recording equipment and hand written notes. Evaluating the DD through field evaluation where users are observed directly as they interact with the DD meant that I did not have to depend solely on the willingness of the participants to complete and return the questionnaires to assess the DD’s usability from users’ perspectives.

While the questionnaire evaluation method did not reveal any new problems not picked up during the field observations, the semi-structured nature of the questionnaire in this study enabled participants to provide qualitative comments regarding the components they found to be problematic during the evaluation sessions.

9.5 COMPARISON OF THE FINDINGS FROM HEURISTIC AND FIELD USABILITY EVALUATIONS

In sections 8.3.2 to 8.3.5, I presented the results obtained from the expert heuristic evaluation of the DD and the analysis of those results. Sections 9.2.2 and 9.2.3 provide the results from the field evaluation and its analysis. In this section, I compare the nature of problems identified by expert evaluators with the actual problems experienced by participants in the field evaluation.

As discussed in section 8.3.2, the output of the heuristic evaluation process was an aggregated list of 71 usability and direct accessibility-related problems. In the field evaluation, a total of 39 problems were encountered by participants. The heuristic evaluation method is known to identify large number of potential usability problems in a relatively cost effective way (see section 4.2.3.1.8). This benefit of the heuristic evaluation method has also been shown in this study.

The heuristic evaluation method was also identified as the most appropriate method to evaluate the direct accessibility support provided in the DD. Two of the five expert evaluators in this study had experience in accessibility evaluation, as further demonstrated by the number of direct accessibility-related problems detected by these two evaluators (see Table 8.5 and Figure 8.10). The detection of direct accessibility-related problems in the DD would have been difficult without the use of the heuristic evaluation method. This is primarily because the DD does not support the use of assistive devices, which made the inclusion of participants with disabilities impossible in the field evaluation. The heuristic evaluation method thus represents the most appropriate method to assess the level of direct accessibility support built into the DD.

The heuristic evaluation method revealed a large number of usability and direct accessibility-related problems, many of which could affect successful interaction with the DD, for example the absence of feedback. Some of these problems could be classified as low severity problems, which may not necessarily affect the use of the system; examples of such include inconsistencies in the use of upper and lower case letters for captions to the desktop icons, and the layout of log in instructions (see Appendix F).

The field evaluation on the other hand revealed actual user problems that impacted on the execution of users' tasks. For example, the configuration of one of the terminals at the school resulted in large icons, which obscured some control buttons and made it impossible for participants to repeat the exercise that had just been completed, as required by the game application following poor performance (section 9.2.3.4). Also, some of the participants could not distinguish between the terms 'username' and 'surname' (section 9.2.3.1).

Twelve of the software usability problems identified during the field evaluation were not detected by expert evaluators. One of these involved a lack of tolerance for user error by the DD. As discussed in section 9.2.3.2, two participants accidentally clicked on the <Cancel> button in the registration form (Figure 8.2). This unintended user error led to the closing of the form without any warning to the participants. None of the expert heuristic evaluators flagged the close proximity of the <Cancel> and <Register User> buttons as potential usability problem. This shows one of the benefits of evaluating with real users, where problems that might be overlooked by expert evaluators are revealed (see section 4.2.3.6.4).

Table 9.3 lists the additional problems experienced by participants during the field evaluation which were not detected by expert evaluators during the heuristic evaluation. Two of the additional problems had to do with the context of use of the DD. As stated in section 9.2.1, the DD used for the observation exercise was installed in an open area, with insufficient shading from the sun. This resulted in reflection from the sun affecting the visibility of pictures in *OpenSpell* (section 9.2.3.5), and the worsening of the already dark background problem in *Themba's Journey* (section 9.2.3.6).

Table 9.3: Additional problems discovered during field usability evaluation

Login screen
1. Confusion between the terms ‘surname’ and ‘username’.
2. An irrelevant rollover that appeared when the pointer is hovered on the username field with the following message “Answer questions here and press Enter when done. For a menu press F10’”.
3. A confusing screen resolution dialogue box with the following message “For best picture quality change the resolution to 1024X. 1: Exit 2: Delete”. This dialogue box appeared momentarily on the login screen and then disappeared after a few seconds
Registration form
4. Contradiction between the lengths of password specified in the hint provided next to the password field and that in the generated error message after the detection of invalid password.
5. Fields with input errors were not highlighted.
6. Close proximity of the <Register User> and <Cancel> buttons.
7. No message for user confirmation following unintended click on <Cancel> button before closing the form.
Main desktop
8. Frequent screen flicker when launching educational game applications.
OpenSpell
9. Unusually large icons due to the configuration of one of the terminals, resulting in control buttons being hidden from viewers.
10. The reflection from the sun affected the visibility of pictures displayed by the application.
Themba’s Journey
11. Very dark background, exacerbated by the open area in which the DD was installed, reduced visibility.
12. Inability to execute the ‘Walk’ option at one of the crossroads in the application.
Keyboard
13. Confusion between the <Enter> key and the ‘mouse click’ keys.
14. Participants did not know that they should use the left arrow key (←) on the keyboard to delete incorrect input in the registration form.

Analysis of the results from the heuristic evaluation, based on the specific interface and application in which the problems were located, revealed that the least number of problems (four) were in the registration form (see section 8.3.4), but the field evaluation results showed that the highest number of problems (eight) encountered by participants was with the form interface. The field evaluation therefore unearthed other usability problems that were overlooked by experts in the registration form.

The findings from the expert heuristic evaluation and the field evaluation with real users showed the benefit and complementary role of combining evaluation methods through expert evaluators with those involving real users, especially in the natural context of the system’s use. As discussed in section 4.2.3, each evaluation method has its benefits and associated limitations. This is one of the reasons why evaluation methods are typically combined in practise [Adebesin, Kotzé and Gelderblom, 2010; Nielsen, 1993; Preece et al., 2007].

The focus of this study was evaluation of a selection of software applications installed on the DD; hence the heuristics were not developed to evaluate the usability of the hardware. The two hardware usability problems discovered during the field evaluation could indicate the need for further studies to assess the usability of DD hardware.

Six of the additional twelve software usability problems uncovered during the field evaluation, and listed in Table 9.3, could have been picked up through the multi-category heuristics by expert evaluators. Table 9.4 lists the six additional problems and the heuristics that are applicable to them. Column 1(problem number) gives the number of the problem as listed in Table 9.3, while column 2 (applicable heuristics) provides the relevant heuristics that could have helped in identifying the problems by expert evaluators. The categories of the heuristics are given in boldface.

Table 9.4: Heuristics applicable to six of the additional usability problems

Problem number	Applicable heuristics
3	**General usability heuristics** 1. Provide information that will enable users understand how to interact with the Digital Doorway using clear and simple terminology. 2. Avoid the use of technical terms.
4	**General usability heuristics** 1. Provide information that will enable users understand how to interact with the Digital Doorway using clear and simple terminology.
5	**Form usability heuristics** 1. When input errors are detected, the cursor should be positioned in the error field with the field highlighted to attract users' attention.
6	**General usability heuristics** 1. Provide support for system exploration by the user by allowing easy reversal of actions. 2. Prevent user error by using appropriate constraints at strategic points.
7	**General usability heuristics** 1. Provide support for system exploration by the user by allowing easy reversal of actions. 2. Prevent user error by using appropriate constraints at strategic points.
11	**Direct accessibility heuristics** 1. Ensure that background and text colours contrast well with each other.

Problems six (close proximity of buttons), seven (no user confirmation) and eleven (dark background) listed in Table 9.3 were serious usability problems that negatively affected users' tasks during the field evaluation. These problems should have been flagged as such by expert evaluators. For example, the close proximity of the <Register User> and <Cancel> buttons (as shown in Figure 8.2) and the very dark background of the game application *Themba's Journey* were quite obvious. Failure of evaluators to detect these problems confirms one of the disadvantages of the heuristic evaluation method (discussed in section 4.2.3.1.9), which is the potential to overlook real user problems.

The reason why problem three (a confusing screen resolution dialogue box with the following message "For best picture quality change the resolution to 1024X. 1: Exit 2: Delete") was not detected is that the dialogue box was not triggered during the evaluation sessions by the expert evaluators. Only two of the nine field study participants encountered the dialogue box.

Table 9.5 lists the remaining six additional usability problems for which none of the multi-category heuristics were applicable. Column 1 gives the number of the problem, relative to its position in Table 9.3 while column 2 provides reasons for the non-applicability of the multi-category heuristics to these problems.

Table 9.5: Additional problems with no applicable heuristics

Problem number	Problem discussion
1	The absence of an applicable heuristic for this problem does not necessarily indicate a weakness in the completeness of the multi-category heuristics. Heuristics are generally effective at revealing predictable usability problems. The confusion that arose between the terms 'surname' and 'username' was as a result of the low level of computer literacy of the study participants. None of the usability principles and guidelines examined in section 4.3 is suitable for revealing this type of problem, which is better uncovered through evaluation with the target user groups for the specific application or system.
2	The appearance of the message "Answer questions here and press Enter when done. For a menu press F10" over the username field on the login screen is questionable. The current keyboard on the DD does not have function keys. It is probable that this code was intended for another application. Adequate testing by the developing team before deployment could have revealed this error. None of the usability principles and guidelines examined in section 4.3 is explicitly suitable for exposing this problem. This type of problem is better uncovered through proper testing of applications during development and prior to deployment.
8	Screen flicker occur as a result of low refresh rate of a computer monitor. This problem could have been classified as a hardware problem. However, it was classified as a usability problem because of the frustration experienced by participants as a result of the need to make multiple clicks before the required application is launched. Although the aim of the study was not to identify hardware usability problems, the general usability heuristic "Response to user action by the system should be instantaneous. Where this is not possible, the system should indicate that the task is in progress to avoid repeated clicking by the user", could be applicable to the user frustration that occurred due to repeated clicks by participants.
9	The configuration of one of the terminals used during the study led to non-visible control buttons. This problem also highlights the problem that could occur as a result of inadequate testing before deployment. It will be difficult to uncover such non-predictable problems through the use of heuristics.
10	Poor visibility of the pictures displayed in *OpenSpell* occurred as a result of insufficient shading from the sun. This was a problem that related to the context of use of the DD, which could not have been detected through the use of heuristics. This highlights the importance of evaluation in a real environment of system usage.
12	The inability of a user to execute the 'Walk' option in *Temba's journey* is likely due to non-implementation of the code associated with this option. This problem also showed that insufficient testing was done before the application was deployed.

	It is noteworthy that none of the general design principles, guidelines and heuristics reviewed in section 4.3 directly addresses the need to ensure that all controls, such as buttons, should be executable. This is perhaps a shortcoming of the design principles and guidelines.

9.6 ASSESSING THE EFFECTIVENESS OF THE MULTI-CATEGORY HEURISTICS

In section 4.2.3.1.3, I discussed the criteria that can be used to assess the effectiveness of the application of a heuristic set in the evaluation of a specific domain, in addition to the number of usability problems identified through the use of the heuristics. These criteria are *correctness* or *terminology*, *coverage* and *thoroughness* [Sim et al., 2009]. This section examines the effectiveness of the multi-category heuristics in the evaluation of the DD

- *Correctness:* This relates to the use of appropriate terminology to describe the heuristics. As discussed in section 7.2.2.2, the heuristics were tested by an HCI expert who had previous experience in usability and accessibility of interactive systems during development. The evaluation process by the expert did not uncover inadequacies in the terms used to describe the heuristics.
- *Coverage and thoroughness:* The primary aim of the evaluation phase of the heuristic derivation cycle, discussed in section 7.2.2, was to assess the completeness of the heuristics. To ensure that the heuristics provide adequate coverage for the applications that will be evaluated, and meet the dual objectives of evaluating the usability and the direct accessibility support provided in the DD, two cycles of evaluation of the heuristics were conducted prior to the formal heuristic evaluation by five expert evaluators. The first evaluation of the heuristics was performed by me while another evaluation was carried out by an HCI expert. Following each cycle, the heuristics were modified based on the findings from these initial evaluation cycles. The multi-category heuristics were also examined in the context of the results from the field evaluation to determine their adequacy in covering actual problems experienced by participants.

  Using the multi-category heuristics, expert evaluators were able to detect large number of potential problems that could affect the usability of the DD. Many of the identified problems were experienced by real users during the field evaluation. Although twelve additional problems experienced by participants were not detected by expert evaluators, six of these could have been detected through the heuristics.

  Non-applicability of the heuristics to the remaining six problems should not be interpreted as weakness of the heuristics, since some of the problems related to the inadequacies in the testing of system functionalities before the DD was deployed.

In addition to the evaluation criteria by Sim et al. [2009], I have also assessed the effectiveness of the multi-category heuristics as discussed below:

- Although Sim et al. [2009] indicate that the raw count of usability problems identified using a heuristic set may not be a good indicator of the effectiveness of the heuristics, the authors were silent on the nature of the usability problems identified as a way of determining the usefulness of the heuristics. The nature of usability and direct accessibility problems identified can be a useful measure of the effectiveness of a heuristic set. As shown in the extract of the problems identified by expert evaluators using the heuristics in Table 8.7 (the complete set of problems identified is provided in Appendix F), many of these problems could negatively impact on user interaction with the DD. Examples of these problems include lack of feedback to users, unintuitive icons and labels, non-adherence to industry standards and convention, and the absence of instructions to users on how to use educational game applications.
- Ling and Salvendy [2005] suggest that the effectiveness of a set of application-specific heuristics can be validated by comparing the results obtained from applying the heuristics in evaluation studies with those obtained by applying a general heuristic set, such as Nielsen's [1994b] heuristics. Rather than applying this method directly, I have modified it by comparing the results obtained from the formal heuristic evaluation with those from the field usability observations.

  As discussed in section 9.5, many of the problems indentified by expert evaluators through the application of the multi-category heuristics were experienced by participants in the field observation. Six of the twelve additional problems that emerged from the field observation could indeed have been picked up through the heuristics, as demonstrated in Table 9.4. The value of the multi-category heuristics was thus made evident through the comparison exercise.

By evaluating the multi-category heuristics used in this study against established criteria, it is easy to demonstrate their applicability to the DD domain. Although the heuristics could not be matched to all user problems in the field study, they were used to discover both major and minor usability and direct accessibility problems in the DD.

9.7 CONCLUSION

In this chapter, I have presented the results obtained from the field usability evaluation conducted to evaluate the DD through real users of the system. The field evaluation was complemented with a semi-structured questionnaire. Of the 37 software usability problems that were encountered by participants in the field study, twelve were not recognized as problems by expert evaluators during the heuristic evaluation sessions. The use of a questionnaire to triangulate the data from the field evaluation also revealed the kind of inconsistencies that may arise between what users say and what they actually do.

To highlight the complementary role of the three evaluation methods, the results obtained from the heuristic evaluation were compared to those from the field evaluation. In addition, I compared the results from the field evaluation with the responses of participants to the questionnaires. I also reflected on the applicability of the multi-category heuristics to the additional problems that were uncovered during the field evaluation.

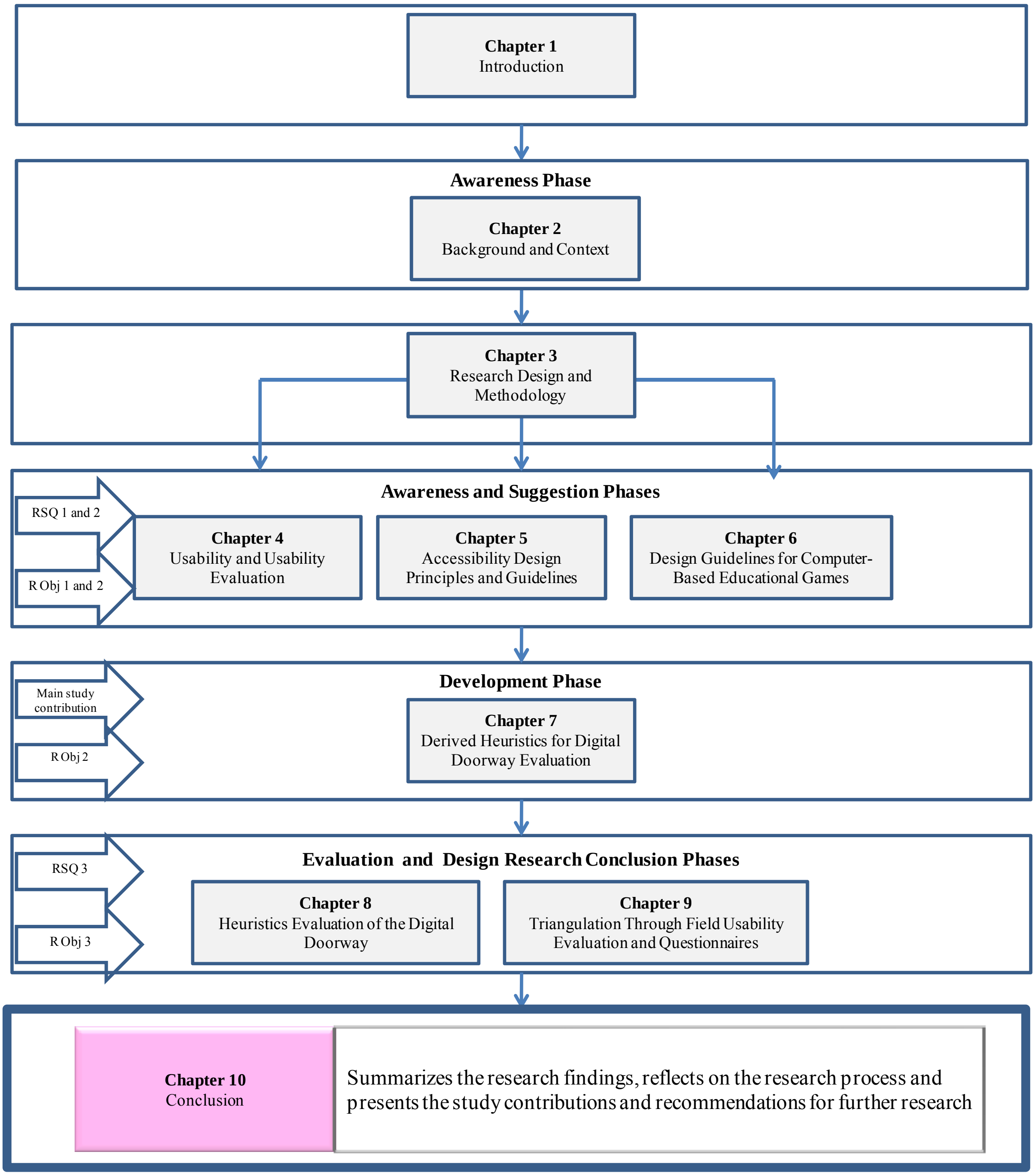
The effectiveness of the multi-category heuristics used by expert evaluators was assessed through:

- The assessment criteria proposed by Sim et al. [2009].
- Consideration of the nature of the usability and direct accessibility problems detected.
- Comparison of results obtained from the heuristic evaluation with those from the field observations.

In Chapter 10, the conclusion to this dissertation, I will provide a summary of the research findings and revisit my research question.

CHAPTER 10: CONCLUSION

The stage of Chapter 10 in the dissertation



Map of Chapter 10

10.1
Introduction

10.2
Summary of Research Findings

10.3
Methodological and Scientific Reflections

10.4
Scientific Contributions

10.5
Secondary Contributions

10.6
Recommendations for Further Research

10.7
Conclusion

10.1 INTRODUCTION

The primary goal of this chapter is to conclude the study and summarize my research findings.

The study involved two cycles of design research phases. The first cycle covered the entire study. At the beginning of this cycle, the need to evaluate the usability and the direct accessibility support provided by DD applications became evident. An investigation of the various methods for evaluating the usability and accessibility of interactive systems was done to determine the methods appropriate for evaluating the DD (sections 4.2.3 and 5.4).

Based on the literature investigation and practical constraints imposed by the DD (see section 3.4.1.1), the heuristic evaluation method was selected as the primary evaluation method. Field evaluation through user observations and questionnaires were employed as complementary and data triangulation methods.

To derive maximum benefits from the heuristic evaluation method, application-specific heuristics are necessary. The need for application-specific heuristics for the DD triggered a second (inner) cycle of the design research. This cycle involved an extensive investigation of existing principles and guidelines for the design of usable and accessible interactive systems (sections 4.3, 5.5 and 6.5). The principles, guidelines and heuristics were examined for their applicability to the DD context, the outcome of which was the set of multi-category heuristics presented in Chapter 7.

Applying the heuristics, expert evaluators were invited to evaluate specific interfaces and applications on the DD, and the results of the evaluation was presented in Chapter 8. The evaluators identified a large number of usability and direct accessibility-related problems and provided suggestions for correcting the problems.

The field evaluation results and the responses of participants to the questionnaire were presented in Chapter 9. The field evaluation revealed twelve additional problems that were overlooked by expert evaluators. Analysis of the returned questionnaires showed some discrepancies between actual behaviour of participants during evaluation sessions and their responses. The combination of the different evaluation methods proved to be beneficial as the weaknesses in one method was offset by the strengths of other methods.

In the next section, I summarize the research findings by revisiting my research question.

10.2 SUMMARY OF RESEARCH FINDINGS

The primary research question for my study was:

> How can standard usability evaluation methods and accessibility evaluation techniques be used in the evaluation of a non-standard system such as the DD?

This primary research question was supported by three research sub-questions:

1. What evaluation methods are available for evaluating the usability and accessibility of interactive systems?
2. Which of these methods can be applied in the evaluation of the DD?
3. What is the result of applying the suitable method(s) to the DD?

This section describes the process that was followed to answer each of the research sub-questions, and then answers the main research question.

Research Sub-Question 1 – What evaluation methods are available for evaluating the usability and accessibility of interactive systems?

During the first cycle of the design research process, I investigated the available methods for evaluating the usability and accessibility of interactive systems (sections 4.2.3 and 5.4). Seven usability evaluation methods were examined in sections 4.2.3.1 to 4.2.3.7. The discussion also included the benefits and limitations of the evaluation methods, and the suitability of each method in formative and/or summative evaluation.

In section 5.4.1, I discussed the accessibility evaluation techniques proposed by Henry [2007]. Henry's [2007] techniques include the heuristic evaluation method, standards review, usability testing, and the use of automatic software tools.

Section 5.4.2 provided the discussion of the lightweight methodology, proposed by Greeff and Kotzé [2009], for accessibility evaluation. This is a three-phase evaluation method involving the use of automatic software evaluation tools, testing with real users, and the establishment of in-house guidelines to guide future development efforts.

Research Sub-Question 2 – Which of these methods can be applied in the evaluation of the DD?

The heuristic evaluation method was found to be a useful and applicable evaluation method for the DD because of its ability to uncover large numbers of usability problems. In addition, because the DD does not currently support the use of assistive devices, the heuristic evaluation method was found to be the most appropriate method to evaluate the direct accessibility support provided in the system.

To obtain maximum benefit from the heuristic evaluation, application-specific heuristics were developed. This was done through an extensive literature investigation of existing usability, accessibility and educational game design principles and guidelines (sections 4.3, 4.4.1.1, 5.5 and 6.5) to determine their applicability to the DD context. The resulting multi-category heuristics (section 7.3) were then used by five independent expert evaluators to evaluate the DD (section 8.3.1).

Because of the potential of the heuristic evaluation method to overlook real user problems, a field usability evaluation using the direct observation method was found to be an appropriate method to evaluate the DD from users' perspectives in a natural environment of DD usage. The field evaluation enabled an assessment of the extent to which the problems identified by expert evaluators were experienced by real users.

To triangulate the data obtained from the field evaluation, the questionnaire evaluation method was also used. A selection of the evaluation heuristics used by expert evaluators was converted into a set of statements that were rated by participants on a five point Likert scale. The questionnaire also provided participants with the opportunity to include comments regarding the problems they experienced during evaluation sessions.

Research Sub-Question 3 – What is the result of applying the suitable method(s) to the DD?

The results of the heuristic evaluation process were presented in Chapter 8. Evaluating the DD through the heuristic evaluation method produced an aggregated list of 71 usability and direct-accessibility problems. The majority of the problems (60.56 percent) were problems that violated general usability heuristics while the least numbers of problems (4.23 percent) had to do with the form usability heuristics. There was a high level of variation in what evaluators considered as being problems, with 48 percent of the aggregated problems being detected by single evaluators.

Using the heuristic evaluation method, the highest number of problems were found in the educational game application, *What-What Mzansi* (seventeen problems) while the least number of problems involved the registration form (four problems). Although the severity ratings for the problems were calculated due to the reasons provided in section 8.3.1, some of the problems could be classified as high-severity problems, for example, there was no feedback following an incorrect username and/or password, and the system did not prevent users from making disastrous mistakes. Other problems, such as, inconsistencies in the use of upper and lower case letters for interface elements, could be rated as having low-severity. The heuristic evaluation process also produced recommendations for correcting the usability and direct accessibility-related problems that were identified.

The field usability evaluation results were presented in Chapter 9. A total of 39 actual usability and direct accessibility-related problems were encountered by participants, two of which were hardware problems. The highest number of problems (eight) occurred while participants were interacting with the registration form, although the heuristic evaluation results showed that the least number of problems were located in this interface. In the field usability evaluation, the lowest number of problems (four) occurred in the educational game, *What-What Mzansi*.

The field usability evaluation revealed twelve additional software usability problems that were not detected by expert evaluators. One of these related to the absence of mechanisms to prevent unintended user error, where two participants accidentally clicked on the <Cancel> button in the registration form. This resulted in the shut down of the form without any warning to the participants. Two of the additional problems related to the context of use of the DD, where insufficient shading from the glaring of the sun affected the visibility of pictures displayed during interaction with

OpenSpell. The dark background of *Themba's Journey* was also made worse by the reflection from the sun.

The return rate of questionnaires by field evaluation participants was low, with only six of the nine participants returning their questionnaires. Of the six participants who returned their questionnaires, only three provided detailed responses. The other three participants did not provide ratings to many of the statements. Moreover, for the few responses that they provided, the ratings were generally positive. The positive responses from these three participants were contradictory to their behaviour during the evaluation sessions. Because the questionnaire was semi-structured, the remaining three participants provided additional feedback regarding the problems they encountered.

Evaluating the DD through the heuristic evaluation method, field usability evaluation and questionnaires complemented one another. Problems that were overlooked by expert evaluators were detected through the field evaluation and some inaccurate ratings to questionnaire statements were clarified through the actual observations that were made during the evaluation sessions.

Having answered the three research sub-questions, the main research question can be addressed.

Primary Research Question – How can standard usability evaluation methods and accessibility evaluation techniques be used in the evaluation of a non-standard system such as the DD?

The primary research question was answered during the process of answering my research sub-questions 1 and 2. Although the evaluation methods discussed in sections 4.2.3 and 5.4 could potentially be applied in the evaluation of non-standard systems, only a few were found to be applicable to the DD context. As discussed in section 4.2.6, the heuristic evaluation method, direct observation in a natural environment and the two query evaluation techniques were found to be appropriate in evaluating the software installed on the DD. With regard to the accessibility evaluation techniques discussed in sections 5.4.1 and 5.4.2, only the heuristic evaluation method was found to be an appropriate method for evaluating the direct accessibility support provided in the DD. However, in the application of the heuristic evaluation method to the DD, the development of application-specific heuristics was required.

The primary research question was thus answered within the context of the DD. It cannot be claimed that this is generally applicable to other non-standard systems. Further research is thus required to confirm the applicability of the evaluation methods employed in this study to other non-standard systems.

10.3 METHODOLOGICAL AND SCIENTIFIC REFLECTIONS

This section provides my reflections on the appropriateness of the chosen research paradigm and the research process.

My study was a typical interpretive research project, which involved data gathering through qualitative research methods. However, the interpretive research paradigm did not adequately match all the activities that were carried out in the study. The study on the other hand fits typical activities of a design research. It involved two cycles of design research phases, with an outer cycle covering the entire research process and an inner cycle involving the development of the multi-category heuristics for the DD.

The process to develop the multi-category heuristics for the DD required an extensive literature investigation of existing usability, accessibility and educational game design principles, guidelines and heuristics. Determining the principles that were applicable and eliminating those that were inapplicable were mainly done by considering the interfaces and applications to be evaluated, the types of users the DD is aimed at and the typical environment of DD usage. This process could easily be judged as being subjective. However, my position is that this was the appropriate method for deriving the application-specific heuristics for evaluating the DD.

As discussed in section 9.6, the multi-category heuristics was assessed as being appropriate for evaluating the DD using the assessment criteria by Sim et al. [2009], a variation of the comparison method by Ling and Salvendy [2005], and by considering the nature of usability and direct accessibility problems identified by expert evaluators. However, this process could have been enhanced by asking the five expert evaluators who used the heuristics, to provide feedback on the completeness and ease of use of the heuristics, following the evaluation sessions.

The severity ratings for the identified problems were not calculated because the evaluation sessions were carried out over three months, which was longer than I had anticipated. My plan was to provide the individual evaluators with an aggregated list of problems identified by all the evaluators and have them provide ratings for the severity of each problem. When it became apparent that the evaluations would take longer than I had planned, I could have asked each evaluator to rate the severity of the problems detected by him/her, rather than wait for the five sessions to be completed. This could have provided some indication of the severity of each of the identified problems.

10.4 SCIENTIFIC CONTRIBUTIONS

According to Kotzé and Johnson [2004], one of the objectives of the field of HCI is to develop or improve productivity and the functionality, safety, effectiveness, and usability of computing systems. Although the usability of interactive systems is important for any user group, it is even more so when the target user population are underserved. Evaluation is an important mechanism for determining the extent to which a given system can be used by the user population to accomplish their goals [Dix et al., 2004]. This study has contributed to the domain of usability and accessibility in the following areas:

- Providing multi-category heuristics for evaluating a non-standard system, such as the DD. The usefulness of the heuristics was tested by five expert evaluators through the heuristic evaluation method.
- Introducing a methodology to develop the multi-category heuristics. As stated in sections 3.4.1.2 and 7.2, the heuristics was developed iteratively through an extensive literature investigation of

multiple design principles and guidelines covering usability, accessibility and educational game applications. This level of rigour was necessary to ensure adequate coverage. The heuristics were further assessed by matching the problems encountered by the field evaluation participants study with the heuristics to determine the extent to which they cover actual user problems.

The use of design research paradigm to guide the development of the multi-category heuristics is also new. While other researchers might have used this paradigm implicitly to develop application-specific heuristics, they did not explicitly state this.

- Extending the eight factors affecting the choice of evaluation methods as identified by Dix et al [2004]. In addition to the factors discussed in section 4.2.4, the nature of the system that will be evaluated also has a major impact on the choice of evaluation method. As discussed in section 3.4.1.1, practical considerations and logistical constraints prevented me from using the usability testing method, even though I had access to, and possess the expertise to use a well equipped usability testing laboratory.

10.5 SECONDARY CONTRIBUTIONS

In addition to the scientific contributions listed in section 10.4, other contributions to the DD project in particular, and similar projects in general, are the following:

- Through the evaluation exercises, this study uncovered potential and real usability and direct accessibility problems in the DD. Many of these problems, for example lack of feedback, unintuitive icons and absence of mechanism to prevent unintended user errors, could have serious implications for the success of the DD project. Without appropriate usability, users may not utilize content that are beneficial.

- In addition to identifying usability and direct accessibility-related problems, the expert evaluators also provided their recommendations for the correction and improvements to the interfaces and applications evaluated (see Appendix F). The correction of as many of these problems as possible will improve the usability of the DD for a wider user group.

- This study also highlights the important role of usability in the efforts to bridge the digital divide. Bridging the digital divide is more that the provision of computing devices. As discussed in section 2.3, to effectively narrow the digital divide, other factors, such as, appropriate usability, accessibility and the relevance of the content provided are crucial [Wilson, 2006].

  During the field usability evaluation, specific problems that related to the environment in which the DD was installed were discovered. Because of inadequacies in the provision of shading from the sun, some of the applications were not visible. This showed that the usability of the DD should be balanced with the need to provide a computer system that is publicly accessible to as many users as possible.

- Although the DD is currently not designed with accessibility in mind, the study has raised the awareness of the DD project team of the need to address accessibility issues in future development of applications for the DD.
- This study has shown the need for the development of in-house usability standards to guide the developers of DD applications. As stated in section 2.4.3, applications that are developed in-house for the DD are implemented by contract and visiting developers. Both the heuristic and field evaluations have revealed the extent of non-adherence to basic usability principles by these developers. One possible reason for non-conformance could be the absence of guiding principles and policies on usability. The set of multi-category evaluation heuristics developed in this study should guide the development of in-house usability standards and guidelines. At the same time, the heuristics could be used to guide the decision in the selection of third-party applications for the DD.

10.6 RECOMMENDATIONS FOR FURTHER RESEARCH

During the course of this study, some further research possibilities were identified. These are:

- Research to look into the applicability of the multi-category heuristics developed in this study to other systems that are similar to the DD.
- A study to evaluate the usability of hardware components of the DD.
- The use of field evaluations with a larger and more diverse user group to determine the extent to which the problems that were identified by expert evaluators, but not experienced by field evaluation participants in this study, are reflected in the use of the DD by other user groups.
- Research to confirm the applicability of the evaluation methods employed in this study to other non-standard systems.

10.7 CONCLUSION

The main purpose of this study was to determine how the standard usability and accessibility evaluation techniques could be used in the evaluation of a non-standard system such as the DD. The heuristic evaluation method, field usability evaluation and questionnaires were found to be the most appropriate methods for evaluating a selection of interfaces and applications installed on DD. The heuristic evaluation method was used as the primary evaluation method, complemented by a field evaluation through user observations and questionnaires.

Application-specific heuristics were developed through an extensive literature investigation of existing principles, guidelines and heuristics for usable and accessible interactive systems for their applicability to the DD context. The main contribution of the study was the set of multi-category evaluation heuristics for the DD using the design research paradigm. Another contribution was an extension of the factors identified by Dix et al. [2004] as affecting the decision to select between different evaluation methods. Other contributions involved the identification of usability problems that could impact on user interaction with the DD, and the recommendations for the improvement of the interfaces and applications evaluated.

REFERENCES

Adebesin, F., Kotzé, 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).*

Adebesin, T. F., De Villiers, M. R. & Ssemugabi, S. (2009). Usability Testing of e-learning: An Approach Incorporating Co-discovery and Think-aloud. *Proceedings of the Annual Conference of the Southern African Computer Lecturers' Association (SACLA 2009).*

Alessi, S. M. & Trollip, S. R. (2001). *Multimedia for Learning: Methods and Development*. Massachusetts: Allyn & Bacon.

Arditi, A. (1999). *Effective Color Contrast: Designing for People with Partial Sight and Color Deficiencies*. Available from: http://www.lighthouse.org/accessibility/effective-color-contrast. Last access date: 05 October 2009.

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. *Universal Access in the Information Society,* 4(3): 270-283.

Bakar, A., Inal, Y. & Cagiltay, K. (2006). Use of Commercial Games for Educational Purposes: Will Today's Teacher Candidates use them in the Future? *Proceedings of the World Conference on Educational Multimedia, Hypermedia and Telecommunications.*

Barnum, C. M. (2002). *Usability Testing and Research*. Massachusetts: Allyn & Bacon.

Bergman, E. & Johnson, C. (1995). *Towards Accessible Human-Computer Interaction*. Available from: http://www.sun.com/accessibility/docs/access_hci.jsp#2.1. Last access date: 30 September 2009.

Berry, B. (2003). Adapting Heuristics for Notification Systems. *Proceedings of the 41st Annual ACM South-East Conference (ACMSE'03).*

Bertot, J. C. (2003). The multiple Dimensions of the Digital Divide: More than the Technology 'Haves' and 'Have Nots'. *Government Information Quarterly,* 20(2003): 185-191.

bridges.org (2006). *Digital Divide*. Available from: http://www.bridges.org/digital_divide. Last access date: 31 May 2010.

bridges.org (n.d). *Real Access / Real Impact Criteria*. Available from: http://www.bridges.org/real_access. Last access date: 31 May 2010.

Brooke, J. (1996). *SUS - A Quick and Dirty Usability Scale*. Available from: http://www.usabilitynet.org/trump/documents/suschapt.doc. Last access date: 28 August 2009.

Burks, M. & Waddell, C. D. (2001). *Universal Design for the Internet*. Available from: http://www.isoc.org/briefings/002/. Last access date: 30 September 2009.

Camacho, K. (2005). Digital Divide. In Ambrosi, A., Peugeot, V. & Pimienta, D. (Eds.) *Word Matters: Multicultural Perspectives on Information Societies.* C & F Editions.

Cambridge, G. (2008). Digital Doorway: Enriching your Mind. *Proceedings of the 2nd CSIR Biennial Conference.*

Card, S. K., Moran, T. P. & Newell, A. (1980). The Keystroke-level Model for User Performance Time with Interactive Systems. *Communications of the ACM,* 23(7): 396-410.

Cedefop (European Centre for the Development of Vocational Training) (2002). *E-learning and Training in Europe*. Available from: http://www2.trainingvillage.gr/download/publication/reference/3021/3021_en_short.pdf. Last access date: 12 January 2010.

Cockton, G., Lavery, D. & Woolrych, A. (2008). Inspection-Based Evaluations. In Sears, A. & Jacko, J. A. (Eds.) *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications.* New York: Lawrence Erlbaum Associates.

Cooper, J. & Kugler, M. B. (2009). The Digital Divide: The Role of Gender in Human-Computer Interaction. In Sears, A. & Jacko, J. A. (Eds.) *Human-Computer Interaction: Designing for Diverse Users and Domains.* Florida: Taylor & Francis Group, LLC.

Corti, K. (2006). *Games-based Learning: A Serious Business Application*. Available from: http://www.pixelearning.com/docs/seriousgamesbusinessapplications.pdf. Last access date: 22 January 2010.

Costabile, M. F. (2001). Usability in the Software Life Cycle. In Chang, S. K. (Ed.) *Handbook of Software Engineering and Knowledge Engineering.* New Jersey: World Scientific Publishing.

Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. California: Sage.

Cruickshank, D. R. & Telfer, R. (1980). Classroom Games and Simulations. *Theory into Practice,* 19(1): 75-80.

Darzentas, J. & Miesenberger, K. (2005). Design for All in Information Technology: A Universal Concern. In Anderson, K. V., Debenham, J. & Wagner, R. (Eds.) *Database and Expert Systems Applications.* Berlin: Springer-Verlag.

David, E. (2008). *Replacing Hearing - Using Captioning*. Available from: http://userlab.usask.ca/ade/replacinghearing.html. Last access date: 19 October 2009.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly,* 13(3): 319-340.

Denis, G. & Jouvelot, P. (2005). Motivation-Driven Educational Game Design: Applying Best Practices to Music Education. *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology.*

Desurvire, H., Caplan, M. & Toth, J. A. (2004). Using Heuristics to Evaluate the Playability of Games. *Proceedings of the Conference on Human Factors in Computing Systems, CHI 2004.*

Dillon, A. (2001). The Evaluation of Software Usability. In Karwowski, W. (Ed.) *Encyclopedia of Human Factors and Ergonomics.* London: Taylor and Francis.

Dix, A., Finlay, J., Abowd, G. D. & Beale, R. (2004). *Human-Computer Interaction*. London: Pearson Education Ltd.

Dondlinger, M. J. (2007). Educational Video Games: A Review of the Literature. *Journal of Applied Educational Technology,* 4(1): 21-31.

Dumas, J. S. (2003). User-Based Evaluation. In Jacko, J. A. & Sears, A. (Eds.) *The Human-Computer Interaction Handbook.* Mahwah: Lawrence Erlbaum Associates.

Dumas, J. S. & Fox, J. E. (2008). Usability Testing: Current Practice and Future Directions. In Sears, A. & Jacko, J. A. (Eds.) *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications.* New York: Lawrence Erlbaum Associates.

Erikson, W., Lee, C. & von Schrader, S. (2010). *United States 2008 Disability Status Report*. Available from: http://www.ilr.cornell.edu/edi/disabilitystatistics/statusreports/2008-pdf/2008-statusreport_us.pdf?cfid=13181917&cftoken=25056954&jsessionid=f0301242e121284efbec27b6367831233726. Last access date: 22 October 2010.

Evers, V. & Hillen, H. (2007). Online Redesign of a Web Site's Information Architecture to Improve Accessibility for Users Who are Blind. In Lazar, J. (Ed.) *Universal Usability: Designing Computer Interfaces for Diverse Users.* Chichester: John Wiley & Sons Ltd.

Fallman, D. (2003). Design-oriented Human—Computer Interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.*

Feng, J. & Sears, A. (2007). Interaction Techniques for Users with Spinal Cord Injuries: A Speech-Based Solution. In Lazar, J. (Ed.) *Universal Usability: Designing Computer Interfaces for Diverse Users.* Chichester: John Wiley & Sons, Ltd.

Fisch, S. M. (2005). Making Educational Computer Games "Educational". *Proceedings of the 2005 Conference on Interaction Design and Children.*

Frishberg, N., Corazza, S., Day, L., Wilcox, S. & Schulmeister, R. (1993). Sign Language Interfaces. *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems.*

Fujitsu (2009). *Web Accessibility Inspector*. Available from: http://www.fujitsu.com/global/accessibility/assistance/wi/. Last access date: 04 January 2010.

Gardner-Bonneau, D. (2010). Is Technology Becoming More Usable -or Less - and With What Consequences. *Journal of Usability Studies,* 5(2): 46-49.

Gelderblom, J. H. (2008). Designing Technology for Young Children: Guidelines Grounded in a Literature Investigation on Child Development and Children's Technology. *Thesis. School of Computing.* UNISA.

Gredler, M. E. (2004). Games and Simulations and their Relationships to Learning. In Jonassen, D. H. (Ed.) *Handbook of Research on Educational Communications and Technology.* New Jersey: Lawrence Erlbaum Associates.

Greeff, M. & Kotzé, P. (2009). A Lightweight Methodology to Improve Web Accessibility. *Proceedings of the Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT 2009).*

Greyling, E. & Smith, R. (2008). An innovative ICT Solution to Steer Rural Communities to Global Understanding: A Case Study From Durban, South Africa. *Proceedings of the World Library and Information Congress: 74th IFLA General Conference and Council.*

Gush, K., Cambridge, G. & Smith, R. (2004). The Digital Doorway - Minimally Invasive Education in Africa. *Proceedings of the ICT in Education Conference.*

Gush, K. & De Villiers, M. R. (2010). Application Usage of Unsupervised Digital Doorway Computer Kiosks in Remote Locations in South Africa. *Proceedings of the Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT 2010).*

Gush, K., De Villiers, M. R., Smith, R. & Cambridge, G. (2010). Digital Doorways. In Steyn, J., Belle, J. P. & Mansilla, E. V. (Eds.) *ICTs for Global Development and Sustainability: Practice and Applications.* Pennsylvania: IGI Global.

Hanson, V. L. (2008). Computing Technologies for Deaf and Hard of Hearing Users. In Sears, A. & Jacko, J. A. (Eds.) *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications.* New York: Lawrence Erlbaum Associates.

Henry, S. L. (2002). Understanding Web Accessibility. In Thatcher, J., Bohman, P., Burks, M. R., Henry, S. L., Regan, B., Swierenga, S., Urban, M. D. & Waddell, C. D. (Eds.) *Constructing Accessibility Web Sites.* Birmingham: Glasshaus.

Henry, S. L. (2007). *Just Ask: Integrating Accessibility Throughout Design.* Available from: http://www.uiaccess.com/justask/. Last access date: 14 August 2009.

Hertzum, M. & Jacobsen, N. E. (2003). The Evaluator Effect: A Chilling Fact About Usability Evaluation Methods. *International Journal of Human-Computer Interaction,* 15(1): 183-204.

Hevner, A. R., March, S. T., Park, J. & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly,* 28(1): 75-105.

HiSoftware (2009). *Cynthia Says*. Available from: http://www.contentquality.com/. Last access date: 04 January 2010.

Hoffman, P. (1999). *Accommodating Color Blindness*. Available from: http://www.stcsig.org/usability/newsletter/9910-color-blindness.html. Last access date: 05 October 2009.

Hornbaek, K. & Frøkjaer, E. (2008). A Study of the Evaluator Effect in Usability Testing. *Human Computer Interaction,* 23(3): 251-277.

Hudson, W. (2004). Inclusive Design: Accessibility Guidelines Only a Part of the Picture. *Interactions,* 11(4): 55-56.

Human Factors Research Group (n.d). *Measuring the Usability of Multi-Media Systems.* Available from: http://www.ucc.ie/hfrg/questionnaires/mumms/info.html. Last access date: 28 August 2009.

IBM (2009). *IBM Software Accessibility Checklist*. Available from: http://www-03.ibm.com/able/guidelines/software/accesssoftware.html. Last access date: 24 November 2009.

IBM (n.d). *Developer Guidelines*. Available from: http://www-03.ibm.com/able/guidelines/index.html. Last access date: 11 March 2009.

International Organization for Standardization (1998). ISO 9241-11 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 11: Guidance on Usability.

Internet World Stats (2010). *Internet World Users by Language*. Available from: http://www.internetworldstats.com/stats7.htm. Last access date: 12 October 2010.

Iwarsson, S. & Ståhl, A. (2003). Accessibility, Usability and Universal Design - Positioning and Definition of Concepts Describing Person-Environment Relationships. *Disability & Rehabilitation,* 25(2): 57-66.

Jacko, J. A., Vitense, H. & Scott, I. U. (2008). Perceptual Impairments: New Advancements Promoting Technological Access. In Sears, A. & Jacko, J. A. (Eds.) *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications.* New York: Lawrence Erlbaum Associates.

Jeffries, R. & Desurvire, H. (1992). Usability Testing vs. Heuristic Evaluation: Was there a Contest? *SIGCHI Bulletin,* 24(4): 39-41.

Jeffries, R., Miller, J. R., Wharton, C. & Uyeda, K. (1991). User Interface Evaluation in the Real World: A Comparison of Four Techniques. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.*

Kahler, H., Kensing, F. & Muller, M. (2000). Methods & Tools: Constructive Interaction and Collaborative Work: Introducing a Method for Testing Collaborative Systems. *Interactions,* 7(3): 27-34.

Keates, S., Kozloski, J. & Varker, P. (2009). Cognitive Impairments, HCI and Daily Living. In Stephanidis, C. (Ed.) *Universal Access in Human-Computer Interaction.* Berlin: Springer-Verlag.

Kennel, A., Perrochon, L. & Darvishi, A. (1996). WAB: World Wide Web Access for Blind and Visually Impaired Computer Users. *ACM SIGCAPH Newsletter,* 55(June): 10-15.

Kieras, D. (2008). Model-Based Evaluation. In Sears, A. & Jacko, J. A. (Eds.) *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications.* New York: Lawrence Erlbaum Associates.

Kjeldsen, R. (2007). An On-Screen Keyboard for Users with Poor Pointer Control. In Stephanidis, C. (Ed.) *Universal Access in Human-Computer Interaction.* Berlin: Springer-Verlag.

Klein, H. K. & Myers, M. D. (1999). A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems. *MIS Quarterly,* 23(1): 67-93.

Korhonen, H. (2010). Comparison of Playtesting and Expert Review Methods in Mobile Game Evaluation. *Proceedings of the 3rd International Conference on Fun and Games.*

Kotzé, P., Eloff, M., Adesina-Ojo, A. & Eloff, J. (2004). Accessible Computer Interaction for People with Disabilities: The Case of Quadriplegics. *Proceedings of the 6th International Conference on Enterprise Information Systems (ICEIS 2004).*

Kotzé, P. & Johnson, C. (2004). *Human Computer Interaction Study Guide for INF120-8*. Pretoria: University of South Africa.

Lazar, J. (2007). Introduction to Universal Usability. In Lazar, J. (Ed.) *Universal Usability: Designing Computer Interfaces for Diverse Users.* Chichester: John Wiley & Sons Ltd.

Lindgaard, G. (1994). *Usability Testing and System Evaluation: A Guide for Designing Useful Computer Systems*. London: Chapman & Hall Computing.

Ling, C. & Salvendy, G. (2005). Extension of Heuristic Evaluation Method: A Review and Reappraisal. *International Journal of Ergonomics and Human Factors,* 27(3): 179-197.

Liu, Y., Osvalder, A. & Dahlman, S. (2005). Exploring User Background Settings in Cognitive Walkthrough Evaluation of Medical Prototype Interfaces: A Case Study. *International Journal of Industrial Ergonomics,* 35(4): 379-390.

Ma, T. H. & Zaphiris, P. (2003). The Usability and Content Accessibility of the E-government in the UK. *Proceedings of the 10th International Conference on Human-Computer Interaction.*

Malone, T. W. (1980). What Makes Things Fun to Learn? Heuristics for Designing Instructional Computer Games. *Proceedings of the 3rd ACM SIGSMALL Symposium and 1st SIGPC Symposium on Small Systems.*

Malone, T. W. (1981). Toward a Theory of Intrinsically Motivating Instruction. *Cognitive Science,* 5(4): 333-369.

Martindale, L. (2002). *Bridging the Digital Divide in South Africa*. Available from: http://www.linuxjournal.com/article/5966. Last access date: 03 May 2010.

Mauri, C., Granollers, T., Lorés, J. & García, M. (2006). Computer Vision Interaction for People with Severe Movement Restrictions. *Human Technology,* 2(1): 38-54.

Mayhew, D. J. (1992). *Principles and Guidelines in Software User Interface Design*. New Jersey: Prentice Hall, Inc.

Millen, D. R. (2000). Rapid Ethnography: Time Deepening Strategies for HCI Field Research. *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (DIS 2000).*

Mitra, S. & Rana, V. (2001). Children and the Internet: Experiments with Minimally Invasive Education in India. *British Journal of Educational Technology,* 32(2): 221-232.

Myers, M. D. (1997). Qualitative Research in Information Systems. *MIS Quarterly,* 21(2): 241-242.

National Institute for The Deaf (2009). *Deaf Affairs*. Available from: http://www.deafnet.co.za/deafaffairs/faq.html. Last access date: 26 October 2009.

National Institute on Deafness and Other Communication Disorders (2002). *Captions for Deaf and Hard-of-Hearing Viewers*. Available from:

http://www.nidcd.nih.gov/health/hearing/caption.asp. Last access date: 19 October 2009.

Newell, A. F., Carmichael, A., Gregor, P., Alm, N. & Waller, A. (2008). Information Technology for Cognitive Support. In Sears, A. & Jacko, J. A. (Eds.) *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications.* New York: Lawrence Erlbaum Associates.

Nielsen, J. (1992). Finding Usability Problems Through Heuristic Evaluation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.*

Nielsen, J. (1993). *Usability Engineering*. California: Academic Press, Inc.

Nielsen, J. (1994a). *Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier*. Available from: http://www.useit.com/papers/guerrilla_hci.html. Last access date: 27 October 2009.

Nielsen, J. (1994b). Heuristic Evaluation. In Nielsen, J. & Mack, R. L. (Eds.) *Usability Inspection Methods.* New York: John Wiley & Sons, Inc.

Nielsen, J. (2003). *Usability 101: Introduction to Usability*. Available from: http://www.useit.com/alertbox/20030825.html. Last access date: 20 July 2009.

Nielsen, J. (2006). *Digital Divide: The Three Stages*. Available from: http://www.useit.com/alertbox/digital-divide.html. Last access date: 30 April 2010.

Norman, D. A. (2001). *The Design of Everyday Things*. London: MIT Press.

Oates, B. J. (2006). *Researching Information Systems and Computing*. London: Sage.

Olivier, M. S. (2004). *Information Technology Research: A Practical Guide for Computer Science and Informatics*. Pretoria: Van Schaik.

Ozok, A. A. (2008). Survey Design and Implementation in HCI. In Sears, A. & Jacko, J. A. (Eds.) *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications.* New York: Lawrence Erlbaum Associates.

Paddison, C. & Englefield, P. (2004). Applying Heuristics to Accessibility Inspections. *Interacting with Computers,* 16(3): 507-521.

Parlangeli, O., Marchigiani, E. & Bagnara, S. (1999). Multimedia Systems in Distance Education: Effects of Usability on Learning. *Interacting with Computers,* 12(1): 37-49.

Pierce, K. (2004). *Using Braille To Access A Computer*. Available from: http://www.esight.org/view.cfm?x=175. Last access date: 15 October 2009.

Pivec, M. (2007). Play and Learn: Potentials of Game-Based Learning. *British Journal of Educational Technology,* 38(3): 387-393.

Preece, J., Rogers, Y. & Sharp, H. (2002). *Interaction Design: Beyond Human-Computer Interaction*. New York: John Wiley & Sons, Inc.

Preece, J., Rogers, Y. & Sharp, H. (2007). *Interaction Design: Beyond Human-Computer Interaction*. Chichester: John Wiley & Sons Ltd.

Pretorius, M. C., Calitz, A. P. & Van Greunen, D. (2005). The Added Value of Eye Tracking in the Usability Evaluation of a Network Management Tool. *Proceedings of the*

Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT 2005).

Pühretmair, F. & Miesenberger, K. (2005). Making sense of Accessibility in IT Design – Usable Accessibility vs. Accessible Usability. *Proceedings of the 16th International Workshop on Database and Expert Systems Applications (DEXA'05).*

Resnik, D. B. (2007). *What is Ethics in Research & Why is It Important?* Available from: http://www.niehs.nih.gov/research/resources/bioethics/whatis.cfm. Last access date: 03 November 2009.

Rigden, C. (1999). 'The Eye of the Beholder' - Designing for Colour-Blind Users. *British Telecommunications Engineering,* 17(Jan.): 2-6.

Ritterfeld, U., Shen, C., Wang, H., Nocera, L. & Wong, W. L. (2009). Multimodality and Interactivity: Connecting Properties of Serious Games with Educational Outcomes. *CyberPsychology & Behaviour,* 12(6): 691-697.

Rogoff, R. (2001). Making Electronic Information Accessible to Everyone. *Proceedings of the Professional Communication Conference, (IPCC 2001).*

Roode, D. (2009). Overview of Research Practice and Research Methodologies. *Seminal Presentation at the CSIR Research and Innovation Core Skills (RICS) Programme*.

Rubin, J. (1994). *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*. New York: John Wiley & Sons, Inc.

Sansonnet, J. P., Braffort, A., Martin, J. C. & Verrecchia, C. (2009). *Towards Interactive Web-Based Signing Avatars*. Available from: http://www.gw2009.de/page6/abstracts/sansonnetea.pdf. Last access date: 28 October 2009.

Sears, A., Young, M. & Feng, J. (2008). Physical Disabilities and Computing Technologies: An Analysis of Impairments. In Sears, A. & Jacko, J. A. (Eds.) *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications.* New York: Lawrence Erlbaum Associates.

Shackel, B. (1991). Usability - Context, Framework, Definition, Design and Evaluation. In Shackel, B. & Richardson, S. (Eds.) *Human Factors for Informatics Usability.* Cambridge: Cambridge University Press.

Shelley, B. (2001). *Guidelines for Developing Successful Games*. Available from: http://jnoodle.com/careertech/files/guidelinesdevelopingsuccessfulgames.pdf. Last access date: 10 March 2010.

Shneiderman, B. (1998). *Designing the User Interface: strategies for Effective Human-Computer Interaction*. California: Addison Wesley Longman, Inc.

Sim, G., Read, J. C. & Cockton, G. (2009). Evidence Based Design of Heuristics for Computer Assisted Assessment. In Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R. O. & Winckler, M. (Eds.) *Human-Computer Interaction - INTERACT 2009.*

Smith, R., Cambridge, G. & Gush, K. (2006). Digital Doorway Computer Literacy through Unassisted Learning in South Africa. *Proceedings of the 1st CSIR Biennial Conference.*

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. & De Villiers, M. R. (2007). A Comparative Study of Two Usability Evaluation Methods Using a Web-Based E-Learning Application. *Proceedings of the Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT 2007).*

Statistics South Africa (2001). *Census 2001: Key Results*. Available from: http://www.statssa.gov.za/census01/html/key%20results_files/key%20results.pdf. Last access date: 03 May 2010.

Statistics South Africa (2007). *Community Survey, 2007*. Available from: http://www.statssa.gov.za/publications/p0301/p0301.pdf. Last access date: 11 August 2010.

Stolterman, E. (2008). The Nature of Design Practice and Implications for Interaction Design Research. *International Journal of Design,* 2(1): 55-65.

Story, M. F., Mueller, J. L. & Mace, R. L. (1998). *The Universal Design File: Designing for People of All Ages and Abilities*. Available from: http://eric.ed.gov/ericdocs/data/ericdocs2sql/content_storage_01/0000019b/80/19/ac/11.pdf. Last access date: 17 November 2009.

Swartout, W. & Van Lent, M. (2003). Making a Game of System Design. *Communications of the ACM,* 46(7): 32-39.

Terre Blanche, M. & Durrheim, K. (2006). Histories of the Present: Social Science Research in Context. In Terre Blanche, M., Durrheim, K. & Painter, D. (Eds.) *Research in Practice: Applied Methods for The Social Sciences.* Cape Town: University of Cape Town Press.

Total Validator (2009). *Total Validator*. Available from: http://www.totalvalidator.com/. Last access date: 08 December 2009.

United Nations (2006). *World Programme of Action Concerning Disabled Persons*. Available from: http://www.un.org/esa/socdev/enable/diswpa01.htm. Last access date: 30 September 2009.

United States Access Board (2000). *Electronic and Information Technology Accessibility Standards (Section 508)*. Available from: http://www.access-board.gov/sec508/standards.htm. Last access date: 09 November 2009.

University of South Africa (2007). *Policy on Research Ethics*. Available from: http://cm.unisa.ac.za/contents/departments/res_policies/docs/researchethicspolicy_apprvcounc_21sept07.pdf. Last access date: 24 July 2009.

Vaishnavi, V. & Kuechler, W. (2004). *Design Research in Information Systems*. Available from: http://desrist.org/design-research-in-information-systems/. Last access date: 17 November 2009.

Vanderheiden, G. C. (1994). *Application Software Design Guidelines: Increasing the Accessibility of Application Software to People with Disabilities and Older Users*.

Available from: http://trace.wisc.edu/docs/software_guidelines/software.htm. Last access date: 04 November 2009.

Verlinden, M., Tijsseling, C. & Frowein, H. (2002). A Signing Avatar on the WWW. In Wachsmuth, I. & Sowa, T. (Eds.) *Gesture and Sign Language in Human-Computer Interaction.* Berlin: Springer-Verlag.

Waddell, C. D. (2002). Overview of Law and Guidelines. In Thatcher, J., Bohman, P., Burks, M. R., Henry, S. L., Regan, B., Swierenga, S., Urban, M. D. & Waddell, C. D. (Eds.) *Constructing Accessible Web Sites.* Birmingham: Glasshaus.

Waraich, A. (2004). Using Narrative as a Motivating Device to Teach Binary Arithmetic and Logic Gates. *Proceedings of the 9th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education.*

Wharton, C., Braffort, J., Jeffries, R. & Franzke, M. (1992). Applying Cognitive Walkthroughs to More Complex User Interfaces: Experiences, Issues, and Recommendations. *Proceedings of the SIGCHI Conference on Human factors in Computing Systems.*

Wildman, D. (1995). Getting the Most from Paired-User Testing. *Interactions,* 2(3): 21-27.

Wilson, C. (1998). *Usability Techniques: Pros and Cons of Co-participation in Usability Studies*. Available from: http://www.stcsig.org/usability/newsletter/9804-coparticipation.html. Last access date: 21 September 2009.

Wilson, C. & Coyne, K. P. (2001). Tracking Usability Issues: To Bug or Not to Bug. *Interactions,* 8(3): 15-19.

Wilson, E. J. (2006). *The Information Revolution and Developing Countries*. London: MIT Press.

World Health Organization (2009). *Causes of Blindness and Visual Impairment*. Available from: http://www.who.int/blindness/causes/en/index.html. Last access date: 06 October 2009.

World Wide Web Consortium (1999). *Web Content Accessibility Guidelines 1.0*. Available from: http://www.w3.org/tr/wai-webcontent/. Last access date: 7 May 2009.

World Wide Web Consortium (2005). *How People With Disabilities Use the Web*. Available from: http://www.w3.org/wai/eo/drafts/pwd-use-web/#deafness. Last access date: 15 October 2009.

World Wide Web Consortium (2008). *Web Content Accessibility Guidelines (WCAG) 2.0*. Available from: http://www.w3.org/tr/wcag20/. Last access date: 5 July 2010.

World Wide Web Consortium (2009). *About W3C*. Available from: http://www.w3.org/consortium/. Last access date: 09 November 2009.

Xerox Corporation (1996). *How to Conduct a Heuristic Evaluation*. Available from: www.stcsig.org/usability/resources/toolkit/pg_he.doc. Last access date: 21 August 2009.

Zimmerman, J., Forlizzi, J. & Evenson, S. (2007). Research Through Design as a Method for Interaction Design Research in HCI. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.*

Appendix A: Formal ethical clearance from UNISA

UNISA university of south africa

Mrs TF Adebesin
P O Box 7216
Pretoria
0001

To whom it may Concern

Permission to conduct Qualitative Research Project

The request for ethical approval for your research project entitled: "Usability and accessibility evaluation of the Digital Doorway using multiple methods" refers.

The College of Science, Engineering and Technology's (CSET) Research and Ethics Committee has considered the relevant parts of the studies relating to the abovementioned research projects and research methodology and is pleased to indicate that the research process poses no ethical problems.

Therefore involved parties may also consider ethics approval as granted. However, the permission granted must not be (mis) construed as constituting an instruction from the CSET Executive or CSET CREC that sampled participants are compelled to take part in the research project. All participants retain their individual right to decide whether to participate or not.

We trust that sampling 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 S Lubbe

Acting Chairperson: CREC

Prof M Setati

Executive Dean: CSET

University of South Africa
Preller Street, Muckleneuk Ridge, City of Tshwane
PO Box 392 UNISA 0003 South Africa
Telephone: +27 12 429 3111 Facsimile: +27 12 429 4150
www.unisa.ac.za

Appendix B: Sample of informed consent form signed by expert evaluators

USABILITY EVALUATION OF THE DIGITAL DOORWAY

EXPERT EVALUATOR CONSENT FORM

Contact: Funmi Adebesin at fadebesin@csir.co.za or 012 841 3373

Dear Sir/Madam

My name is Funmi Adebesin. I am a Masters Studentship holder from the Meraka Institute of the CSIR and studying at University of South Africa (UNISA). My research involves evaluating the usability of the Digital Doorway.

This is a request for you to participate as an expert evaluator of the Digital Doorway. Results from this evaluation will be used for research purposes only and findings may be included in articles in research journals or papers in conference proceedings; however, the identity of participants will be protected with no specific reference to them by name.

Yours sincerely

Funmi Adebesin

I, ____

hereby voluntarily agree to participate as an expert evaluator of the Digital Doorway.

Signature: ______________________ Date: ______________

Appendix C: Sample of formal letter of request to school principal

USABILITY EVALUATION OF THE DIGITAL DOORWAY

Contact: Funmi Adebesin at fadebesin@csir.co.za or 012 841 3373

PERMISSION TO CONDUCT RESEARCH AT GATANG SECONDARY SCHOOL

Dear Sir/Madam

My name is Funmi Adebesin. I am a Masters Studentship holder from the Meraka Institute of the CSIR and studying at University of South Africa (UNISA). My research involves evaluating the usability of the Digital Doorway, that is, how easy the computer is to use. As part of my research, I wish to conduct field observations at your school where I will observe learners using the Digital Doorway. After using the Digital Doorway, learners will also complete a questionnaire where they will assess the Digital Doorway.

The purpose of the study is to evaluate the Digital Doorway and **not** the ability of the learners. Information collected from learners will assist in improving future versions of the Digital Doorway.

Permission for learners to participate will also be obtained from their parents/guardians. All information collected will be treated as **private** and **confidential** and the identity of the learners will not be revealed. Participation of learners in this study is **absolutely voluntary** and they are free to withdraw their participation even after their parents have signed the consent form

I will be grateful if you would give me your permission to conduct the study at your school.

Should you wish to obtain further information regarding the evaluation please do not hesitate to contact me at the above address.

Yours sincerely

Funmi Adebesin

I, ____

the principal of, ______________________________________

hereby consent to your performing your research at my school.

Signature: ______________________ Date: ______________

Appendix D: Sample of informed consent form signed by parents/guardians of field evaluation participants

meraka INSTITUTE

CSIR our future through science

USABILITY EVALUATION OF THE DIGITAL DOORWAY

Contact: Funmi Adebesin at fadebesin@csir.co.za or 012 841 3373

PERMISSION FOR LEARNER TO PARTICIPATE IN RESEARCH

Dear Sir/Madam

My name is Funmi Adebesin. I am a Masters Studentship holder from the Meraka Institute of the CSIR and studying at University of South Africa (UNISA). My research involves evaluating the usability of the Digital Doorway, that is, how easy the computer is to use. As part of my research, I wish to conduct field observations at Gatang secondary school where I will observe learners using the Digital Doorway. After using the Digital Doorway, learners will also complete a questionnaire where they will assess the Digital Doorway.

The purpose of the study is to evaluate the Digital Doorway and **not** the ability of the learners. Information collected from learners will assist in improving future versions of the Digital Doorway.

Permission to conduct the research at Gatang has also been obtained from the principal. All information collected will be treated as **private** and **confidential** and the identity of the learners will not be revealed. Participation of learners in this study is **absolutely voluntary** and they are free to withdraw their participation even after you have signed this consent form.

I will be grateful if you would give permission for your child to take part in the study.

Should you wish to obtain further information regarding your child's participation in the evaluation please do not hesitate to contact me at the above address.

Yours sincerely

Funmi Adebesin

I, ____

the parent/guardian of ________________________________

hereby permit him/her to participate in your research.

Signature: ______________________ Date: ______________

Appendix E: Information document for expert evaluators

Researcher: Funmi Adebesin fadebesin@csir.co.za

Dear evaluator,

Thank you for taking part as an expert evaluator for the Digital Doorway.

- Please take time to familiarize yourself with the following documents to be used in the evaluation process. The documents are:
  1. *DD_HE_Information*, Page 1 of this document.
  2. *DD_HE_System-Overview*, Page 2 – 4 of this document.
  3. *DD_HE_User-Profile*, Page 5 of this document.
  4. *DD_HE_Eval-Proc*, Page 6 of this document.
  5. *DD_HE_Expert-Profile*, a separate document.
  6. *DD_HE_Eval-Criteria*, a separate document.
  7. *DD_HE_Evaluation_Report_Sheet*, a separate document.
- Please provide the information required in the document, *DD_HE_Expert-Profile*, and e-mail it to me at: fadebesin@csir.co.za.
- Please let me know when you will be available to perform the evaluation on the Digital Doorway at the Meraka Institute of the CSIR between 15 September 2009 and 30 November 2009; the evaluation should take about two hours.
- I will be present during the evaluation to answer any question you may have regarding the evaluation of the Digital Doorway.
- You are free to make notes of the evaluation, otherwise I could facilitate the session by acting as the scribe. This will allow you to concentrate on the most important thing – identifying usability/accessibility problems in the Digital Doorway. After the evaluation, I will compile a report of usability/accessibility problems identified which will be sent to you the on the day following evaluation at the latest. Please verify this report to ensure that it is a true reflection of the evaluation session.
- Please e-mail the verified report together with any comment to me at: fadebesin@csir.co.za.
- Should you require further information regarding the evaluation, please contact me at the above e-mail address or call 012-841-3373. Once again, thank you for evaluating the Digital Doorway.

Application to be evaluated: The Digital Doorway

The Digital Doorway is a non-standard, non-Web based computer system developed as a joint initiative between the South African Government's Department of Science and Technology (DST) and Meraka Institute of the Council for Scientific and Industrial Research (CSIR). The computer is housed in a rugged, custom-designed kiosk with three terminals to curtail vandalism. It has metal keyboards with metal trackballs that enable simultaneous access by three users. One of the terminals acts as a server, which runs on the Ubuntu Linux operating system. The other two terminals are diskless clients, which access files from the file server. Some of the programs and applications in the Digital Doorway were developed in-house but a large number of them are third-party resources and programs.

The Digital Doorway is based on the 'Hole in the Wall' concept from India, where Mitra and Rana [2001] demonstrated people's innate ability to acquire basic computing skills through experimentation and self-discovery provided that the technology is made available in an environment conducive to experimentation.

The Digital Doorway is depicted below in Figure 1:

(Note: The picture of the Digital Doorway is the same as that in Figure 2.2 and has been deleted from this sample document to prevent duplication)

The evaluation will be limited to those interfaces and applications developed in-house by the Digital Doorway team. The specific interfaces and applications that will be evaluated Digital Doorway login screen, the form interface for creating a new user account, the educational quiz game, *What-What Mzansi*, the educational spelling game *OpenSpell* and the life-skills game, *Themba's journey*. Accessibility evaluation will be restricted to the direct accessibility features built into the Digital Doorway.

Overview of Educational games that will be evaluated

What-What Mzansi

What-What Mzansi is quiz game in the form of yes/no questions. The interface includes three menu items: <Play>, <Hi-Scores> and <About>

Play: As the name suggests, this is where the questions are asked and answered. There are two levels of difficulty from which a player can select: 'Easy' and 'Advanced'. On the selection of a level, a local voice welcomes the player and reads out the question which can be answered by clicking on 'Yes' or 'No'. Each question lasts 60 seconds; the score for each question can ranges from 2 to 10, depending on how fast the player is able to answer it; for instance, if the player clicked on yes when the score mark is on 6, then his/her score will be 6 if the answer is correct. However, if the answer is wrong, the score value will be deducted from the current total score. At the end of the session, all the questions are presented again with the correct answers. If the player has performed well, he is informed of his readiness to move to the next level; otherwise he can repeat the session.

Hi-Scores: This submenu lists all the high scores registered for a number of users so far for the game.

About: This menu option mainly contains information on the Digital Doorway project and its achievements.

What-What Mzansi is only available in English.

Open Spell

Open spell is a program designed for the teaching of spellings. It is available in all the eleven South African official languages from which the user can select. There are three difficulty levels to choose from.

The interface presents three menu options <Say>, <Guess> and <Spell>. It also includes letters of the alphabet, similar to a keyboard and three control buttons <Repeat>, <Erase>, and <Enter>.

Say: Selecting this menu option brings up a picture of the word to be spelt one after the other. A voice in the selected language speaks out each letter of the word as well as its pronunciation.

Guess: This functionality is based on the hangman word guessing game. The specific word to be guessed is presented as dashes which represent the number of alphabets in the word. The player selects a letter of the alphabet from the screen by clicking on it. If the chosen letter appears in the word it is slotted in the appropriate dashed space(s). With each incorrect guess, a bird is perched on a tree branch accompanied by an audio effect of the branch breaking-off the tree. This process is continued until the player guesses the correct word or the tree branch breaks off and the birds fall down.

Spell: When this option is selected, a picture is presented and the name of that picture is read out. The player is then prompted to spell the word.

Themba's Journey

This is a life skills game that tells the story of Themba, who makes a journey from his village to the city. At strategic places, Themba reaches crossroads where the user has to make decisions on his behalf on a course of action, with each action having a positive or negative consequence. The game is available in both IsiXhosa and English languages.

The main interface includes three menu items, <Help>, <Play> and <Exit>.

The *Play* menu is where the main story is narrated.

The *Help* menu contains navigation instructions and instructions on how to play the game.

The *Exit* menu enables users to close the program.

Profile of Digital Doorway Users

The Digital Doorway is targeted at adult and young users with little or no computer literacy. However, children and young adults are currently the main user groups. Digital Doorways are typically being used by single and multiple user groups. The number of those using the system as a group number can vary from 4 to 12 users.

Digital doorways are typically installed at schools, libraries and community centres.

Procedure for conducting the heuristic evaluation on the Digital Doorway

1. Log-in as a guest user by following the instruction on the screen.
2. On the desktop, search for the educational games, *What-What Mzansi*, *Themba's Journey*, and *Open Spell*.
3. Thereafter, open the new_content folder to access the educational games, *What-What Mzansi*, *Themba's Journey*, and *Open Spell* in whatever order you wish and familiarize yourself with the games.
4. Log out of the Digital Doorway.
5. Examine the log-in interface, taking note of all associated usability/accessibility problems and the heuristics violated.
6. Identified problems should be related to the scribe who will make notes (when using a scribe). Please be specific when describing problems and their locations.
7. Now, create a new user account by completing the registration form.
8. This form is one of the interfaces being evaluated; using the set of heuristics provided, take note of or relate all the associated usability/accessibility problems in the registration form as well as the applicable heuristics that were violated.
9. Log in using the user-name and password you have chosen.
10. Now re-access *What-What Mzansi*, *Themba's Journey*, and *Open Spell* one after the other in whatever order you wish and evaluate each game using the set of heuristics provided.
11. While evaluating each game in turn, take note of or relate all usability/accessibility problems in each game as well as the heuristics that has been violated to the scribe.
12. For each problem identified, please provide recommendation(s)/solution(s) to the problem, if you can.
13. If a scribe was used during the evaluation, the scribe will compile a report of all usability/accessibility problems identified after the evaluation, The evaluation report will then be mailed to you a day after the evaluation at the latest. Please review the report to ascertain that it reflects the evaluation process, adding comments as required. The reviewed report should then be mailed to: fadebesin@csir.co.za.
14. If you have made note of the evaluation session yourself, please send your detailed report on the evaluation using the template evaluation report sheet provided. Mail the report to: fadebesin@csir.co.za as soon as possible.

Please accept my sincere appreciation for taking part in evaluation the Digital Doorway.

Appendix F: Aggregate usability and accessibility problems identified by expert evaluators

Column 1: Category of the heuristics as presented in sections 7.2.2.1 to 7.2.2.4.

Column 2: The number of the heuristic relative to its category as presented in sections 7.2.2.1 to 7.2.2.4.

Column 3: The description of specific usability/accessibility problems.

Column 4: The evaluator(s) who identified the problems.

Column 5: Evaluators' recommendations for correcting the problems.

Heuristic category	Heuristics violated	Usability/Accessibility problems	Identified by (Evaluator)	Recommendations for Improvement
Login Screen				
1	1.1	The benefit/use of creating a new user account as opposed to using the system as a guest is not clear. The login instruction is quite confusing, not sure how to handle the choice between creating a new user account and using the system as a guest.	E2; E4	1. Provide information that will assist users in making the decision of either to create a user account or use the system as a guest user. This is particularly important in the context of Internet access via GPRS using recharge vouchers
1	1.1; 2.3	Instructions for creating user account and for guest login are lumped together in the same text box.	E1	1. Separate the two instructions
1	1.2	After entering username, there is no indication of what to do next.	E1; E2; E3	1. Provide an <Enter> button next to username and password textboxes that can be clicked by the user. 2. Provide a prompt for the user to press the Enter button on the keyboard.
1	2.4	No option available to retrieve forgotten password	E2	1. The standard for system requiring username/password login is to have a mechanism for retrieving forgotten password
1	4.1	There is no feedback whatsoever when an incorrect username and/or password is provided.	E1; E2; E3; E5	1. Give explicit feedback that will enable users realize what their mistake.
1	2.1	After the acceptance of username and password, the message 'Language en_ZA.UTF-8 does not exist. Using system default" was displayed.	E1	1. Use simple terms comprehensible to novice user.

3	1.1	The font sizes of the instructions on how to login/create user account in four other languages are small.	All	1. Provide login instruction in the predominant language for the specific area of deployment. 2. Consider an intelligent interface screen which will default to the preferred language of the user (selected at registration) following successful login using username/password 3. It would be good to have a language option selection button, and then have the instruction displayed in the preferred language 4. Increase the font size of these instructions
3	3.2	Background is colour too dark.	E5	1. Change to a brighter colour.
New Account Registration Form				
1 2	1.2 3.1; 3.2	There is no indication of which fields must be filled and which ones are optional.	E2	1. Clearly specify required and optional fields
1	2.1	The instruction/prompt '6 – 10 characters' in the hint next to username and password data fields are technical and may not be understood by novice users.	E1; E2	1. Rather say '6 to 10 letters or numbers'
2	1.1	At the start of form application, the cursor is not positioned in the first data field. The user is required to place the cursor in the first field	E1; E3; E5	1. Provide a visible flashing cursor when the form is loaded in the first data field.
2 3	1.2 4.1	User cannot use the <Tab> key on the keyboard to select female for the gender field	E3	1. Allow the use of tab keys to move around form fields.
Positive remarks regarding the registration form 1. The provision of sample data expected for personal details is good. 2. The separation of personal details from user information is good.				
Main Desktop				
1	1.1; 2.2	It is difficult to determine the differences between the functionalities of the icons with the captions 'chat' and 'kchat'.	E1; E3; E4	1. Use descriptive name that will enable users determine the functionality of icons.
1	1.1; 2.2; 3.1	The functionality of the right pointing arrow button ⇒ is not clear from its look This button is also hidden from users' view.	All	1. Provide a clearly marked exit button whose functionality is easily discernible to user. 2. Use rollovers to provide simple information on its functionality.

1	1.1; 2.2; 3.1	The functionality of the volume control slider is not clear from its look. Information in the volume control dialogue window is highly technical.	E1; E5	1. Rather use the conventional speaker icon which can be adjusted by sliding the bar up and down 2. Use rollovers to provide simple information on its functionality.
1	2.2; 3.1	The same icon/graphic is used to represent different objects on the desktop.	E1; E2; E3	1. Use graphics that will be suggestive of its functionality to users. 2. Use roll-overs which are displayed when the mouse hovers around icons to provide an indication of their functionality.
1	2.2; 3.1; 9.1	One of the elements on the desktop has the label/caption **'?'**. This is not descriptive of its function.	E1; E2; E3; E4	1. Use descriptive name that will enable users determine the functionality of icons.
1	2.2; 9.1	One folder on the desktop has the caption 'new_content'. This is not descriptive of the applications it contains	All	1. Use descriptive name that will enable users determine the functionality of icons.
1	2.4	The <System> button, which include the <logout> button, is too far away from the <Programs> and <Resources> menu options.	E3	1. Move the <System> button to the left of the screen.
1	2.2; 3.2	The game applications *What-What Mzansi*, *Themba's journey* and *OpenSpell* are hidden inside the folder 'new_content'.	E1; E2; E3; E4	1. Place the game applications What-What Mzansi, Themba's journey and Open Spell directly on the desktop to facilitate high visibility. 2. Include these three games within the Programs→Games submenu since other game applications are located here.
1	5.1; 5.2	The location of the following icons on the taskbar are too close: ⇒, <System> <Volume control>, and Volume control slider. Users can easily click on the ⇒ button while trying to use the volume control slider, thereby closing the system unintentionally.	E5	1. Separate the volume control icons from the system exit icons.
1	8.1	The caption of the element labelled 'Bluetooth_ saver' has the first character capitalized while the captions for all other icons and folders are in lowercase.	E1	1. Be consistent in the use of uppercase and lowercase when labelling icons/objects.
1	8.2; 8.3	There is lack of consistency in clicking effect. Some elements require double clicking to activate (for example the new_content folder) while others can be activated by a single click (for example, What-What	E4	1. Be consistent in the ways system functionalities can be activated

<table>
<tr><td></td><td></td><td>Mzansi).</td><td></td><td></td></tr>
<tr><td>1</td><td>9.1; 9.2</td><td>The label for the folder ‘new_content’ gives the impression that users will be presented with new information every time it is activated. In fact, there in nothing ‘new’ about the contents.</td><td>E4</td><td></td></tr>
<tr><td>3</td><td>3.2</td><td>The level of contrast between the dark blue background and the grey foreground used to label icons is low.
The contrast between the word ‘Digital Doorway’ and the dark blue background is poor</td><td>E3; E5</td><td>1. Change to a brighter colour.</td></tr>
<tr><td colspan="5">Other general comments on the Digital Doorway and content of main desktop
1. While there are excellent contents under <Programs> and <Resources>, however, the use of these resources is not promoted due to poor visibility. I had to discover these contents while trying to figure out how to exit the Digital Doorway.
2. There is lack of interactivity in the tutorials located in the icon labelled ‘?’
3. The tutorials are not learner paced.
4. No audio presentations for tutorials.
5. Some of the tutorials include elapsed time slide, which provides an indication how long the tutorial will take. However, other tutorials do not have the time slide.
6. Although the desktop interface is simple, the aesthetics could be improved.
7. The keyboard is quite robust and appropriate for environment of use.</td></tr>
<tr><td colspan="5">What-What Mzansi</td></tr>
<tr><td>1</td><td>2.2</td><td>The functionality of the icon ‘?’, which provides context-specific instructions about the game, might not be adequately interpreted by some users.</td><td>E5</td><td>1. Provide a rollover indicating its function.</td></tr>
<tr><td>1</td><td>2.4; 2.5</td><td>The icon ‘>’ within <About> menu option is typically used for forward progression, however, clicking on it takes the user back to the main screen of the game.</td><td>E5</td><td>1. Change this icon to the backward progression symbol ‘<’ and provide a rollover to indicate this functionality.
2. Provide an option for users to go straight to play the game (If they so wish), rather than having to go back to the main screen.</td></tr>
<tr><td>1</td><td>4.1</td><td>The marks provided in “Total Score” and “Score” are different. This is confusing.</td><td>E4</td><td></td></tr>
<tr><td>1
4</td><td>1.1
2.1</td><td>The character reading out instructions and questions tells the user to answer yes or know, but did not tell him/her how to do this.</td><td>E3</td><td>1. The spoken instruction should also include a prompt to answer by clicking on either of the pots with the words ‘yes’ and ‘no’.</td></tr>
</table>

1 4	3.2 2.3	The countdown timer is not highly visible at its current location. This makes it difficult for users to know they are being timed.	E5	1. The timer is an important part of this game so it should be located at more prominent space to make it visible to users.
1 4	3.2 6.2	At the start of the application, some of the control buttons and the character that reads out instructions and questions are hidden from user's view. A full screen mode of the application is activated by clicking arbitrarily around the taskbar.	All	1. The full screen view should be automatically displayed at the start of the application
1 4	9.1; 9.2 2.1	The label <About> is misleading. Evaluators were expecting to find game instructions here.	E2; E4; E5	1. Provide clear instructions regarding the purpose of the game, how to navigate, rules, constraints. 2. While the current content of is good, it should ideally be located at the Digital Doorway home page
3	4.1	The <PgDn> and <PgUp> buttons on the keyboard cannot be used to navigate the content of the window within the <About> menu option.	E5	1. Provide for keyboard navigation. 2. Provide rollovers for the symbols for paging up and down.
3 4	5.3 6.2	Volume control not visible when game is viewed in full screen mode.	E1	1. Provide a local volume control within the game in addition to the global volume control
4	2.1; 2.3	Learner demotivation can easily occur because despite high score, one cannot proceed to the next level. Progress actually depends on the number of correct questions. User get heavily penalised if incorrect answer is provided very fast (for example, if slider is on 10 when you answer, you get -10 for incorrect answer).	E4; E5	1. Provide explicit rules governing the game so users are aware of what is expected of them.
4	4.2	The performance feedback 'don't make me laugh' after a poor performance is cheeky and not encouraging. Some users might find it offensive	All	1. Provide feedback in ways that are not demeaning.
4	5.1 5.2	Incorrect answers cannot be modified by users. Performance feedback does not provide the correct answers to wrong answers chosen by users.	E2; E4	1. In order to provide effective learning, the feedback should explicitly provide the correct answer rather than just indicating that user-provided answer is wrong.
4	6.1	Frequent users should be able to skip introductory music	E2	1. Provide an option for user to go straight to playing game.

4	6.1; 6.4	There is no <Exit> or <Forward> button available when the <About> menu option is selected	E3; E4; E5	1. Provide an explicit <Exit> and <Back> buttons inside <About> window.
4	6.2	After reading the content in <About> menu option, it was difficult to proceed further as the control button was hidden (because of earlier problem with full screen view)	E4; E5	
4	6.3	One cannot exit the game when in full screen view because the <X> button is not functioning. User must change to around 70% screen view and exit with the browser close button. The <X> button, which presumably should close the game window, is not working.	All	1. Correct the associated code that implements this button. 2. Provide an <exit> button option on the main menu
4	7.1	The game is only available in English	E1	1. Consider making the game available in other languages if possible.
Positive remark regarding What-What Mzansi 1. The game is highly interactive. It is enjoyable.				
OpenSpell				
1	2.2	The use of the labels <Say>, <Guess>, and <Spell> are not descriptive of their functionalities.	All	1. Use a more descriptive term
1	2.2; 3.1	The * symbols used to represent level of difficulty are not intuitive.	All	1. Use descriptive terms which will enable users understand the functionality. 2. Separate the language selection option from level of difficulty as two menu options.
1	2.4	The onscreen keyboard does not follow standard keyboard layout.	E2	1. Modify onscreen keyboard to follow standard keyboard layout.
1	4.2	When a correct answer is provided in <Guess> mode, the screen disappears very fast with no performance feedback.	E4; E5	1. Provide feedback that persists long enough for user observation. 2. Consider providing clapping sound effects to acknowledge correct answer.
1	9.1	The menu labelled <Game> is being used for language selection options while another one which is appropriately labelled <Language> (suggesting that a user can select choice language) is greyed out.	All	1. Change the label of <Game> menu to <Language>. 2. Remove the greyed out items if the functions are not available for users to select

1 4	1.1 2.1	There is no instruction on how to play the game and use the buttons.	All	1. Provide clear and easily accessible instructions on how to play and use the control buttons
1 4	9.1 2.1	The caption on the menu <About> is misleading. Was expecting to find instructions about the game here.	E4; E5	1. While the current content of is good, it should ideally be located at the Digital Doorway home page
3	1.1	The font size for menu palette at the top-left of the application window is small and somewhat hidden from the user's view	E3	1. Increase the font size of the menu labels a little.
3	1.5	The quality of the speech output is poor and not easily discernible even when the volume is at maximum level.	All	1. Improve the quality of the speech output
3	2.1	No visual feedback for correct/wrong answers in <Spell> mode.	E2	1. In addition to providing verbal feedback, also provide a thick mark to indicate correct answer and an X mark to indicate wrong answer
3	5.1; 5.2	The speed at which words are spelt by the voice 'instructor' is fast. A user cannot review previous spellings.	E5	1. Provide mechanism that will enable the user repeat previous spellings.
1 3 4	10.1; 10.2 4.1 2.3	When the <Spell> option is selected, the user cannot use the keyboard to provide input but must use onscreen keyboard.	E1; E2; E3; E4	1. Allow users to provide input via the keyboard in addition to using onscreen keyboard.
4	4.3; 5.2	There is no corrective feedback when wrong answer is given while playing in the <Guess> mode.	E5	1. Provide corrective feedback that will enable users learn from mistakes.
4	6.1	The user cannot control which words are spelt while using the <Say> option. The user cannot use the <Repeat> button if he wants the last spelt word repeated.	E1; E3; E4	
Themba's Journey				
1	2.4	When the mouse pointer is hovered on the speech bubbles, it changes to a hand with pointing finger. This is typically used to indicate a 'clickable' object	E3	1. Change to pointer with arrow head.
1	8.1; 8.3	Themba's journey provides an <Exit> button to close application, while the same functionality is provided	E5	1. Be consistent in the way icons proving similar functionalities are represented

		by the <X> button in What-What Mzansi.		
1	8.3	Language choice selection is via a menu in OpenSpell; however, the same action requires the user to hover the mouse over speech bubbles in Themba's journey.	E1; E2; E3; E4	
1 4	1.2; 7.2 2.1	At each of the crossroads, there are no instructions on how make selection between available choices. Although information regarding this is provided under <Help> many users would have forgotten when it is actually needed.	E4; E5	1. Although information regarding the crossroads is provided in <Help>, users would have forgotten about the information when it is actually required. Rather provide a simple instruction on the same screen (for example, a prompting question asking "what should Themba do?").
1 4	5.1; 5.2 7.5	Some users might accidentally click on the <Exit> button.	E5	1. Provide a dialogue box where users can confirm whether they actually want to exit the program.
1 4	8.2 7.1	The road safety information and other life skills tips are provided in English; however, the story line had consistently been given in IsiXhosa.	E2; E3; E4; E5	1. Be consistent in the way things are done. 2. Provide these information in IsiXhosa as well.
3	1.4	Road safety information and other life skills tips is only available in text.	E2	1. Provide speech equivalent of the information in addition to the text.
3 4	1.2 7.1	To access an English version, the user must hover the mouse on the speech bubble. This is problematic for users with limited use of their hands. The instruction provided under <Help> did not specify this.	All	1. Provide an English version of the program in speech in addition to text version. 2. Provide a menu option from where the user can easily select language of choice.
3 4	1.4 7.1	There is no audio equivalent for instruction in <Help>. The instruction in <Help> is provided only in English, not IsiXhosa, the default language.	E5	1. Make the instruction available in the default language as well.
3 4	1.4 7.1	The narration voice is only in IsiXhosa. Non-Xhosa users who cannot read cannot use the application.	E2; E3; E4; E5	1. Provide an English version of the program in speech in addition to text version
4	1.1	Purpose of life skills tips might be missed by some users.	E5	1. It might be worthwhile having another persona voice making explicit the importance of adhering to road safety rules. For instance, this can be done having the voice say something like "let's go over the lesson we have learnt from the previous scene".

4	3.3	Long story. Require some time to complete.	E5	1. Provide a mechanism which keeps track of where the user temporarily stopped the application so he/she is not compelled to start all over again (at least for registered users).
4	6.2	The backward/forward buttons disappears at the crossroads, making it difficult for user to review previous screens.	E2	1. Make all control buttons visible and accessible in all screens
4	6.3	The main <Exit> button does not close the application. An <Exit> button is available on all the scenes of the application but clicking on it takes the user back to the main window.	All	1. Correct the associated code that implements this button.
4	6.4	There is no <Exit> button on the <Help> menu screen.	E1	1. Provide an explicit <Exit> button within the <Help> window.
Positive comments regarding Themba's Journey 1. The highlighting on the control buttons on being selected is good as this visibly show users that a button has been selected.				

Appendix G: Peer reviewed papers generated from study

1. Adebesin, A; Kotzé, P; Gelderblom, H. 2010. The Complemetary Role of Two Evaluation Methods in the Usability and Accessibility Evaluation of a Non-Standard System. Proceedings of the SAICSIT 2010 Conference – Fountains of Computing Research, ACM Press, p. 1 – 11, ISBN: 978-1-60558-950-3.

   The paper received the "Best Research Paper by a Masters Student" award.

2. Adebesin, A; Kotzé, P; Gelderblom, H. 2010. The Impact of Usability on Efforts to Bridge the Digital Divide. Proceedings of the 4th IDIA 2010 Conference – Exploring Success and Failure in Development Informatics: Innovation, Research and Practice: Monash University, ISBN: 978-0-620-47590-7.

Copies of the papers are provided in the following pages.

The Complementary Role of Two Evaluation Methods in the Usability and Accessibility Evaluation of a Non-Standard System

Funmi Adebesin
Meraka Institute of the CSIR and School of Computing, UNISA
P.O. Box 395
Pretoria, 0001, South Africa
+27 12 841 3373
fadebesin@csir.co.za

Paula Kotzé
Meraka Institute of the CSIR and School of Computing, UNISA
P.O. Box 395
Pretoria, 0001, South Africa
+27 12 841 4791
paula.kotze@meraka.org.za

Helene Gelderblom
School of Computing, UNISA
P.O. Box 392, UNISA, 0003, South Africa
+27 12 429 6631
geldejh@unisa.ac.za

ABSTRACT

Usability, which is generally defined in terms of application effectiveness, efficiency and user satisfaction, is one of the focus areas of human-computer interaction (HCI). Accessibility is the design of systems that can be perceived, understood and used by people with varying abilities. Although accessibility concerns are aimed at making systems usable for people with disabilities, support for direct accessibility, the built-in redundancies in an application that enable as many people as possible to utilize it without system modifications, is beneficial to people with or without disabilities. Different usability evaluation methods (UEMs) are available. Selecting between the various methods can be influenced by the type of system being evaluated. The Digital Doorway (DD), a non-standard computer system deployed to promote computer literacy amongst underprivileged communities in South Africa, was evaluated using the heuristic evaluation method and a field usability study. The heuristic evaluation method revealed a large number of usability and direct accessibility-related problems, some of which could be classified as low-severity problems. The field study showed additional problems that affected the successful completion of user tasks. Since a number of these were a direct consequence of the context of use, they were not recognized as problems by expert evaluators. The study showed that the heuristic evaluation method can be optimized by complementing it with another method that involves user participation and is, preferably, carried out in the intended context of use.

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *evaluation/methodology, User-centered design.*

General Terms

Design, Human Factors.

Keywords

Accessibility, field evaluation, heuristic evaluation, usability.

1. INTRODUCTION

Usability, one of the focus areas of human-computer interaction (HCI), 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" [18].

The success of any interactive system is dependent on, amongst other factors, its utility, ease of use from the user's perspective, and user experience [5; 6; 28]. It is therefore crucial that designers incorporate usability design principles early on in the design process, and evaluate systems for usability before deployment. When the target user group has special needs this becomes even more important.

Accessibility used to be a concern of the built environment practitioners where legislations required cuts into street curbs to enable easy access by people in wheelchairs. Now, the explosive growth of the World Wide Web (WWW) and the Internet has put pressure on designers to make electronic information accessible to people with disabilities [19; 32]. Accessibility in the context of HCI is defined as the design of applications that "are perceivable, operable and understandable for people with a wide range of abilities" [14].

Some authors consider accessibility to be a subset of usability [13]. Others [31] claim that it is a prerequisite for usability. However, what is incontrovertible is that both design concepts contribute to good design. The specified users in the definition of usability for any given application may be diverse, including those with disabilities. Incorporating usability and accessibility design guidelines will result in an application that is usable by a wider spectrum of users.

The primary focus of this paper is on usability evaluation although we also address the issue of accessibility evaluation by considering the direct accessibility properties of the system being evaluated. Direct accessibility refers to the built-in redundancies in applications which enable as many people as possible to utilize it without system modifications [39]. Direct accessibility can enhance general usability, offering benefits to people with or without disabilities.

Different approaches to usability evaluation exists, among them heuristic evaluation, field observation, surveys and controlled usability testing. Some of these methods require expert evaluators while others rely on end-user involvement [6; 26;

30]. Several previous studies compared usability evaluation methods (UEMs), for example, those by Gray and Saltzman [10] and Molich and Dumas [25].

Combinations of usability evaluation methods may be required to offset the limitations inherent in the use of any single evaluation method, for example, the heuristic walkthrough evaluation method [34]. The focus of this paper is on the complementary role that the combination of expert and user-based evaluation methods can play, especially when the target system is a non-standard system[1].

The Digital Doorway (DD) is an example of such a system. It is a walk-up and use system deployed amongst underprivileged communities around South Africa as part of the effort to narrow the digital divide. Ever since the installation of the first DD in 2002, the DD project has mainly focused on providing physical computer access to underprivileged communities around the country. The systems were deployed without any formal usability evaluation of the software applications installed on them.

DDs are housed in rugged, custom-designed kiosks with multiple terminals that can be accessed simultaneously by users. Each of the terminals has a metal keyboard with reinforced touchpad for input [12]. The robust housing and metal keyboard help to minimize vandalism. Pre-loaded software applications and contents run on the Ubuntu Linux operating system, however, the interface does not follow any particular design standard or operating system interface. The systems are installed at venues such as schools, police stations and community centers. The project has seen 206 DDs installed around the country since its inception. A three-terminal DD is shown in Figure 1.

Figure 1. A three-terminal Digital Doorway

http://www.digitaldoorway.org.za/index_main.php?do=hardware

Despite the fact that the project has focused on hardware development, the hardware does not currently support the use of assistive devices, such as screen readers for visually impaired users. Furthermore, the environment of use of the DD sometimes impose additional restrictions on the use of the system e.g. noise and glaring of the sun. It has thus become essential for us to also consider the level of direct accessibility support in the evaluation of the DD, in addition to evaluating the usability of software installed on the system.

This paper describes the use of multi-category heuristics or evaluation criteria (both terms used interchangeably) for evaluating a non-standard computer system and how the method was complemented with a field usability evaluation that retained the system's context of use. The results of the two methods were compared to underscore their complementary role.

The rest of the paper is structured as follows: Section 2 briefly describes some of the UEMs, highlighting their benefits and limitations. In Section 3, we discuss how the heuristic evaluation method was applied in the evaluation of the DD and the results obtained from the evaluation. Section 4 describes the field usability study on the DD. In Section 5, we compare the results obtained from the two studies and conclude the paper in Section 6.

2. OVERVIEW OF USABILITY EVALUATION METHODS

This section provides a general overview of UEMs, highlighting the benefits and limitations of each.

UEMs are generally classified into two main groups – expert analysis and user evaluation. The following subsections briefly describe the heuristic method, cognitive walkthrough, direct field observations, interviews, questionnaires, and controlled usability testing.

2.1 Heuristic Evaluation

The heuristic evaluation technique, pioneered by Nielsen and Molich in 1990 [27], involves expert evaluators independently critiquing an interface using a set of evaluation criteria in order to identify potential usability problems. The heuristic method is an easy, flexible and cost effective method that can be used in formative and summative evaluations. However, to be of any value, the set of evaluation criteria must be appropriate for the specific application and at least three expert evaluators are required [20; 27; 30].

2.2 Cognitive Walkthrough

This is an inspection method that is based on cognitive science theory where experts step through a set of tasks. Cognitive walkthrough aims to assess the learnability of systems where the preferred method of learning to use the system is by exploration rather than going through user manuals [4; 6; 30; 40]. It is a flexible method that can be used for formative and summative evaluation before user testing [40], but it does not address other measures of usability like the application efficiency. The method assumes that the evaluator possesses cognitive theory skills [40].

2.3 Direct Field Observation

This involves the evaluator observing the users as they carry out normal or routine activities in the natural context of use either at home or the workplace. Because the context is retained, direct field observation can reveal details that are difficult to obtain using other evaluation methods. This natural environment may sometimes disrupt the evaluation process as a result of high level of noise and constant interruptions from colleagues. In addition, participants may alter their behavior when they become aware of being observed [7].

2.4 Interviews and Questionnaires

These are query techniques that can be used to elicit users' requirements for a proposed system or measure the extent to which an implemented system meets their expectations. Interviews can be structured, semi-structured or a combination of both. Interviews, especially unstructured interviews, are

[1] Non-standard in this context means systems that do not display standard operating system interfaces or use standard equipment.

beneficial since the questions can be varied to allow the evaluator to probe issues as they arise to obtain deeper understanding [21; 30]. Interviews can be highly subjective as they are typically used to measure users' opinion. While it may be impossible to avoid participant subjectivity, it is essential to be aware of them.

Questionnaires are similar to interviews in that the questions can be closed or open, though they are not as flexible because the questions are fixed and further probing is often impossible. While questionnaires can be used to reach large number of respondents in less time compared to interviews, their distribution and return rates can be problematic. The potential for generating flawed data is high when using questionnaire as the sole evaluation method. Hence it may not adequately reflect the actual usability of the application [26]. This is why evaluation methods are typically combined to triangulate data [26; 30].

2.5 Controlled Usability Testing

Usability testing (UT) is an evaluation method where the performance of typical users is measured as they carry out real, pre-defined tasks using the target application. The aim of UT is to test the usability of the system, not the users' ability. UT is expensive, requiring sophisticated usability laboratory equipped with monitoring cameras and equipment [3; 30; 33]. To be effective, UT is typically combined with think-aloud, where users are encouraged to verbalize their thoughts and the reasoning behind their actions as they carry out the tasks [3].

Table 1 provides a summary of the evaluation methods discussed above, highlighting the advantages and limitations of each method.

Table 1. Summary of usability evaluation methods

Technique	Description	Advantages	Limitations
Heuristic evaluation	Experts independently assess the interface using a set of evaluation criteria.	- Flexible, can be used for formative or summative evaluation. - Ability to provide quick feedback to designers. - Can reveal large numbers of potential usability problems	- Sometimes require the development of application-specific evaluation criteria. - Requires multiple evaluators. - Some problems that may affect real users may be overlooked.
Cognitive walkthrough	Experts step through a set of tasks to assess learnability of the system.	- Flexible, can be used in formative and summative evaluation.	- Setting up of representative tasks can be tedious. - Other measures of usability are not addressed.
Direct field observation	Users are observed while carrying out tasks in a natural context of use.	- Natural context is retained. - Ability to reveal specific usability problems which may impact user tasks.	- High level of disruptions. - Participants may alter their behavior.
Interviews	Used to obtain user requirements for a new system and measure extent to which a functional system meet their expectations.	- Flexible, can be used for formative or summative evaluation	- Participants' subjectivity.
Questionnaires	Used to obtain user requirements for a new system and measure extent to which a functional system meet their expectations.	- Large number of respondents can be reached in short time.	- Participants' subjectivity. - Low return rate.
Usability testing	Real users are observed while carrying out pre-specified tasks in a controlled environment.	- Ability to reveal specific usability problems which may impact user tasks.	- Expensive, requires sophisticated equipment. - Some participants may find thinking aloud unnatural.

2.6 Choice of Evaluation Methods

The decision to select between the different evaluation methods is dependent on a number of factors, including the stage in the development life cycle at which evaluation is done (formative or summative), whether evaluation should be conducted in a controlled environment or natural setting and availability of resources [6]. For example, the heuristic evaluation method is flexible in that it can be used for formative or summative evaluation, provided appropriate evaluation criteria are used.

In this study, the main factor that influenced our choice of evaluation methods is the type of system being evaluated. Although we have access to a well-equipped usability testing laboratory, practical considerations make the use of controlled usability testing unfeasible since we could not physically move the DD to the usability laboratory. Furthermore, observation and logging software in the usability laboratory is only compatible with the Windows operating system. As stated in section 1, applications on the DD run on the Ubuntu Linux operating system. Because the study involved the evaluation of a fully functional system, the heuristic method is appropriate. To complement the heuristic evaluation method, a field usability evaluation was also done at a local school where the DD is installed.

3. HEURISTIC EVALUATION OF THE DIGITAL DOORWAY

In this section, we describe the application of the heuristic evaluation method in the evaluation of the DD.

The DD provides access to a large number of software applications and other resources, the majority of which are open source or third-party applications. These includes the OpenOffice suites, educational games, scientific simulations, Wikipedia documents and Mindset applications, the latter being a South African curriculum-based educational program [11]. The content can be accessed by logging in as a guest (using the username 'dd1' for example) or as a registered user. A new user account is created by completing a simple electronic form, which is activated by typing 'new' in the username field.

Because it is impractical to evaluate all the applications on the DD, a selection of interfaces and applications developed in-house were evaluated in this study. These are: the login screen, the new user account registration form, the main desktop, and three educational games – *What-What Mzansi* (a quiz game), *OpenSpell* (an educational spelling game), and *Themba's Journey* (for developing life-skills).

3.1 Developing the Heuristics

As stated in Sections 2.1 and 2.6, the heuristic evaluation method can be used in formative or summative evaluation provided the evaluation criteria are suitable for the system being evaluated. To come up with evaluation criteria that provide adequate coverage of interfaces and applications to be evaluated, we used contextual analysis to examine the principles for usable interface design by Dix, Finlay, Abowd, and Beale [6], Gelderblom's [9] guidelines for the design of children' technology, Mayhew's guidelines for the design of form-fill interfaces [24], Nielsen's heuristics [27], the usability principles by Preece, Rogers and Sharp [30], the design principles by Norman [29], and Shneiderman's [36] golden rules for interface design, for their applicability to the DD system.

To address direct accessibility concerns, we studied the seven universal design principles by Story, Mueller and Mace [37], the web content accessibility guidelines (WCAG 1.0) by world wide web consortium (W3C) [41] (WCAG 2.0 was still in draft form when the heuristics were derived), the United States' standards for electronic information accessibility (Section 508) [38], and the IBM software accessibility checklists [17].

Because the applications evaluated included educational games, there was a need to study guidelines relating specifically to such applications. Game-specific guidelines studied are those proposed by Alessi and Trollip [2], Malone [22; 23] and Shelley [35].

Evaluation criteria were derived systematically over three iterations. During the first iteration, the first author of this paper tested the generated heuristics on the DD to assess their adequacy. The heuristics were then modified in another round of iteration and tested by the third author on the DD again. Further modifications were then made before producing a final set of evaluation criteria for the DD.

Not all the principles and guidelines examined were equally applicable to the DD. Examples of non-applicable principles are those relating to multithreading, task migratability and the use of markups and style sheets [22].

The process of refinement yielded a total of 77 heuristics. Page restrictions do not permit the inclusion of the complete heuristic set, but the interested reader can access it in [1]. A representative subset of the heuristics is provided in Table 2.

To aid readability and facilitate the analysis of identified problems, the criteria were organized into four categories, namely general usability, form-filling, direct accessibility, and game heuristics.

Table 2. Subset of evaluation criteria for the Digital Doorway

Category 1: General Usability Heuristics	
1.1	Provide information that will enable users understand how to interact with the DD using clear and simple terminologies.
1.2	Provide clear indication of what the next required action is.
1.3	Icons, labels and symbols should be intuitive and meaningful to users, taking into account user context and experience.
1.4	Follow and adhere to platform and industry standards and conventions.
1.5	Be consistent in the naming conventions used for icons, symbols, and objects.
1.6	Objects, options, and permissible actions should be visible so that users do not have to remember instructions.
1.7	Audio instructions should be given close to when the user is expected to act on them.
1.8	Feedback should be provided in clear and unambiguous terms.
1.9	Response to user action by the system should be instantaneous. Where this is not possible, the system should indicate that the task is in progress to avoid repeated clicking by user.
1.10	Prevent user error by using appropriate constraints at strategic points.
1.11	Error messages should be context-specific in relation to the action performed.
1.12	Ensure that information sequence and layout appear in natural and logical order.
1.13	The DD should not impose unnecessary constraints on user input method.
Category 2: Form-Filling Heuristics	
2.1	Provide visible cue by positioning the cursor in the first data field at start of the form.
2.2	Ensure that related items are grouped together to aid readability.
2.3	Provide visual reinforcement for element groups through efficient use of white spaces and borders.
2.4	Designate required fields in standard and consistent ways taking into account users' age and experience.
2.5	When input errors are detected, the cursor should be positioned in the error field with the field highlighted to attract user's attention.
Category 3: Heuristics to support direct accessibility	
3.1	Text size of instructions should be large enough to enable easy perception by users with low vision.
3.2	Provide audio equivalent of instructions and information to afford access by users who cannot read.
3.3	Provide quality speech output that enable users hear and comprehend their meanings.
3.4	Provide feedback using multiple modes to facilitate access and comprehension.

3.5	Ensure that colour alone is not used to represent important information.
3.6	Ensure that background and text colours contrast well with each other.
3.7	Allow keyboard navigation for operations/tasks that do not essentially require use of the mouse.
3.8	Provide controls that enable users to pause, continue, or repeat audio-visual information.
3.9	Information should be accessible without undue physical efforts.
Category 4: Game-specific Heuristics	
4.1	Games should have clear goals and objectives.
4.2	Provide an easily accessible instruction on how to play the game.
4.3	Permissible actions and constraints should be clearly specified.
4.4	Users should be able to adjust the game's level of difficulty.
4.5	Performance feedback should not be given using negative or sarcastic statements.
4.6	Provide constructive and corrective feedback that will enable player learn from mistakes and improve future performance.
4.7	All control mechanisms should be visible and easily accessible.
4.8	Provide clear exit route to enable users leave the game at any stage.
4.9	Game should be accessible in different languages.

3.2 The Evaluation Process

The DD was voluntarily and independently evaluated by five usability/accessibility experts using the generated heuristics, part of which is provided in Table 2. Three of the experts have experience in usability evaluation while the other two have expertise in usability and accessibility evaluation. The evaluators were provided with the evaluation criteria, the procedure to be followed, and an overview of interfaces and applications to be evaluated, well in advance so they could familiarize themselves with the relevant documentation. Evaluators also signed informed consent forms that guarantee their anonymity and the confidentiality of information they provided.

For practical reasons (discussed in Section 2.6), evaluators had to go physically to the DD laboratory located at the CSIR for the evaluation. The five evaluation sessions were conducted over a 3-month period.

Using Nielsen's [27] procedure for conducting heuristic evaluation with some modifications the evaluators went through a two-pass session, with the first pass aimed at getting a feel of how the applications work. In the second pass they did the evaluation. Each session lasted approximately 2 hours although we did not set any time limit. The sessions were facilitated by one of the authors who acted as scribe.

Following the evaluation, an evaluation report was compiled by the scribe and mailed to the relevant evaluator within two days of the evaluation. The report was then verified by the evaluator to ascertain that it is a true reflection of the evaluation. In some cases, the verification resulted in modifications to the report.

Taking notes and compiling evaluation reports allowed us to make judicious use of evaluators' time and reduced the associated difficulty of note taking and evaluation while standing. Thus the evaluators were able to focus on the evaluation process.

3.3 Evaluation Results

Between the five evaluators, an aggregate of 71 usability and/or accessibility problems were identified. Thirty-four problems were identified by single evaluators (47.88% of the total problem set) while ten (14.08%) were recognized as such by two evaluators. Seven (9.85%) problems were identified by three evaluators, and a further seven (9.85%) were identified by four evaluators. All the five evaluators agreed that 13 (18.30%) of the total problem set were problems.

Within the login screen, a total of eight problems were identified by evaluators, four were located in the new user account registration form, and 13 on the main desktop. For the educational game applications, 17 problems were located in *What-What Mzansi*, 14 in *OpenSpell* and 15 in *Themba's Journey*. The number of problems located in each interface and application is shown in Figure 2.

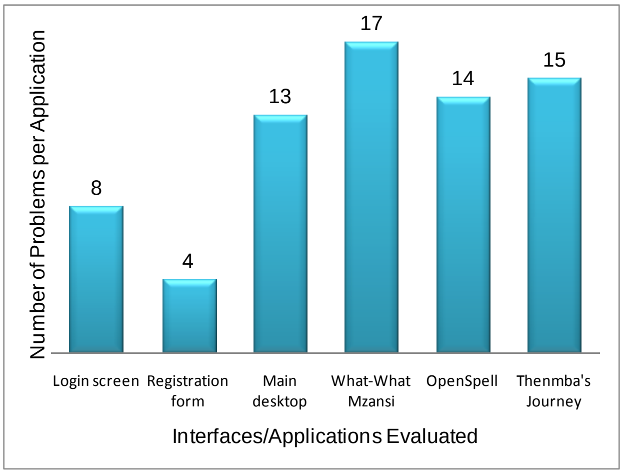


Figure 2. Number of problems identified in each interface and application

Due to page restrictions, the complete list of problem descriptions (available in [1]) could not be presented here. In section 5 we provide a subset of identified problems and compare them with actual problems experienced by users during field evaluation.

4. FIELD USABILITY EVALUATION OF THE DIGITAL DOORWAY

Designers typically make assumptions about potential users [8]. Evaluation with real users enables evaluators to assess the extent to which those assumptions are valid.

In this section, we discuss the field usability evaluation conducted to evaluate the DD in one of the centers where it is deployed.

4.1 The Evaluation Environment

We evaluated the DD at a local school where the system is installed in an open area on one of the school's corridors. This provides unrestricted access to the system. Children from surrounding homes also have access to the DD soon after the school closes until 18:00 late in the afternoons.

Prior to conducting the study, a formal approval was obtained from the school principal. Parents/guardians of participants also signed informed consent forms.

Six learners participated in the evaluation, with two participants each using one of the educational games *What-What Mzansi*, *OpenSpell* or *Themba's Journey*. Participants were also required to register a new user account before accessing the applications unless they had a valid account. Participants with disabilities were not included in the evaluation, since the DD does not support the use of assistive devices such as a screen reader. Hence, it will be pointless to include a blind user, for example.

4.2 The Evaluation Process

Rather than use the conventional field study where users are observed while using the system, participants were given pre-defined tasks (shown in Figure 3) to complete. This enabled us to focus the evaluation on the specific interfaces and applications identified for the evaluation.

Digital Doorway Evaluation -Task list

1. Read the screen instruction on how to register as a new Digital Doorway user if you are not a registered user.
2. Complete the registration form if you are not a registered user, otherwise proceed to step 3.
3. Start the Digital Doorway by providing the requested information.
4. Search for the quiz game '**What-What Mzansi'**.
5. Remember to provide verbal feedback all the time.
6. Search for and read the instruction on how to play the game.
7. Proceed to play the quiz game.
8. Choose how difficult you want the game to be.
9. Change the volume to suit your need.
10. Remember to provide verbal feedback all the time.
11. Close the Digital Doorway when you are done.

Figure 3. Task list for field usability evaluation (using the application 'What-What Mzansi)

To avoid disruptions to learning activities as much as possible, and minimize distractions from noise, evaluation sessions took place in the afternoons well after the official closing hour of the school.

Using the cooperative evaluation style, participants were encouraged to ask questions and assistance whenever they got stuck with any activity. This approach is justified since the DD is not a transaction processing system where the speed of task completion is a measure of usability.

As a secondary instrument to triangulate data, each participant was given a semi-structured questionnaire to evaluate the DD after the evaluation sessions. The questions in the questionnaires were based on a selection of the derived heuristics, discussed in section 3, although they were rephrased using simpler terminology to aid comprehension by novices.

4.3 Evaluation Results

Various usability and direct accessibility-related problems were encountered by users during the evaluation. These ranged from inability to locate the required application, to lost data due to lack of error tolerance by the system.

The total number of actual user problems experienced by the six participants was 29, as illustrated in Figure 4. Eight of these were additional problems not recognized by expert evaluators while the others formed part of the problems identified by the experts. Three problems were related to the login screen, six were in the new user account registration form while five involved the main desktop. In the educational game applications, three problems related to the quiz game, *What-What Mzansi*, six problems involved *OpenSpell* while a further six related to *Themba's Journey*.

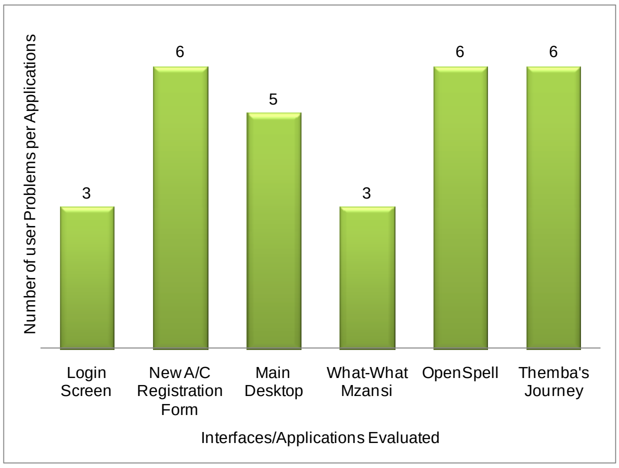


Figure 4. Number of actual user problems per interface/application

Only 4 of the 6 questionnaires were returned. Analysis of the questionnaires showed discrepancies between participants' responses and their actual behavior during observations. For example, in response to the statement, *I am able to determine the meaning and purpose of signs and symbols used in the Digital Doorway*, two participants responded with 'Agree'. However, one of these participants was unable to locate the volume control and exit buttons on the desktop without assistance while the other could not set the level of difficulty in OpenSpell (designated with *s).

Although the questionnaire method was not a primary method of evaluation in the study, the responses from participants highlighted some of the problems associated with the use of questionnaires as mentioned in Table 1. In Section 5, we provide the usability and direct accessibility-related problems found during the field evaluation and compare them with those identified by expert evaluators.

5. COMPARISON OF RESULTS FROM HEURISTIC AND FIELD USABILITY EVALUATIONS

As discussed in section 2, each evaluation method has benefits and limitations. Combining an expert evaluation method like heuristic evaluation with one that involves user participation, especially in a natural environment of use, allows one method to offset the shortcomings in the other, for example by revealing problems that were not picked up while using the other method.

Table 3 provides the description of the problems identified by expert evaluators and compare them with actual user problems during field evaluation to emphasize this complementary role.

Table 3. Problems identified by expert evaluators versus actual user problems

Heuristics violated	Problems identified by expert evaluators	Actual user problems
Login Screen		
1.8	There was no feedback whatsoever when an incorrect username and/or password was provided.	Incorrect username and/or password were a common user error. The system returned the same screen over and over without an indication of what the mistake was. Most of the time the field observation facilitator had to inform participants that the problem was with the username or password they are providing.
1.1	The login instruction was quite confusing, not sure how to handle the choice between creating a new user account and using the system as a guest.	The system did not provide login information for users who have just registered or those with existing accounts. The only prompt on the screen is addressed to guest users and those wanting to create new user accounts. Some participants actually typed in ‘new’ or ‘dd1’ to login, while others asked what should be done next before being told that the newly chosen username and password should be used to log in.
1.1, 1.12	Instructions for creating new user account and that for a guest log in are lumped together in the same text box.	None.
1.3	None	Users sometimes confused their surname with a ‘username’ and type their surname in the username field, though this was not the chosen username.
3.1	The font sizes of the instructions on how to login/create user account in four other languages, namely Xhosa, Afrikaans, Sotho and Venda were small.	None. Participants in the field evaluation only read the English version of screen instructions. The font size of instructions in the other languages could have been a problem if the participants did not understand English.
1.2	After entering the username, there was no indication of what to do next.	User problems related to this involved the hardware i.e. the keyboard. Some participants confused the <Enter> key on the keyboard with the key designated for producing a ‘mouse click’ effect because the keys were not labelled. However, after pressing one without the desired effect they then pressed the other.
New User Account Registration Form		
2.1	When the form application is activated, the cursor is not positioned in the first data field. The user is required to place the cursor in the first field	Some participants began typing their names only to realise later that the input was not being accepted and needed to place the insertion point within the first field before typing again.
2.4	There is no indication of which fields are compulsory to be filled and which ones are optional	Participants typically kept the home language and preferred language fields empty only to have error messages urging them to fill the fields.
3.7	User cannot use the <Tab> key on the keyboard to select female for the gender field	None of the participants used the <Tab> key to navigate the form. They positioned the insertion point over relevant fields before typing or clicking to make their selections. This was to be expected since they are not expert users.
1.10	None	Two participants accidentally clicked on the <Cancel> button. This inadvertent user error resulted in the form being closed without any warning to the user thereby erasing all the data fields input thus far.
2.5	None	The form did not facilitate the location of an error field. A participant erased his input in the password field accidentally, while trying to correct the name field entry following an error message. The insertion point remained in the password field after clicking on the <Register user> button. Without the user realising this, he pressed the backspace button several times and erased the wrong field.
Main Desktop		
1.1, 1.3	The functionality of the volume control slider is not clear from its look.	Only three of the six participants were able to locate the volume control button on the desktop, the other three required assistance after several failed attempts.

1.3	One folder on the desktop has the caption 'new_content'. This is not descriptive of the applications it contains. The game applications *What-What Mzansi*, *Themba's journey* and *OpenSpell* are hidden inside the folder 'new_content'.	Only two of the participants found the location of the game applications on their own. Other participants unsuccessfully searched for the applications within the <Game> submenu, located in the Resource menu, before they were told where to find them.
3.6	The level of contrast between the dark blue background and the grey foreground used to label icons is low. The contrast between the word 'Digital Doorway' and the dark blue background is poor.	Three of the participants found the background colour to be too dark. On several occasions, they had to shield their faces and the screen with their hands while using the DD to overcome the extent of reflection of the sun on the dark background. The dark background was significantly worse than that experienced in the close-up laboratory used by expert evaluators. The reflection worsened the contrast issue.
1.5	The caption of the element labeled 'Bluetooth_ saver' has the first character capitalized while the captions for all other icons and folders are in lowercase.	None
1.10	The locations of the following icons on the taskbar are too close to one another: the right-pointing arrow button ⇒, <System> <Volume control>, and Volume control slider. Users could easily click on the right-pointing ⇒ button while trying to use the volume control slider, thereby closing the system unintentionally.	A user accidentally clicked on the ⇒ button used to exit the system while trying to locate the volume control button and the system was shut down without any warning.
	What-What Mzansi	
1.6	At the start of the application, some of the control buttons and the character that reads out instructions and questions are hidden from user's view. A full screen mode is activated by clicking on an icon which does not indicate this function.	None of the two participants who used this application knew how to get a full screen view of the game.
1.6	None	One of the terminals (the third terminal) used for the evaluation sessions had unusually large icons. This resulted in non-visibility of a number of control buttons, in this particular instance, a right pointing arrow '>' used for forward progression. This made it impossible for participants to repeat the level which they had just completed as required following poor performance.
4.5	The performance feedback 'don't make me laugh' after a poor performance is cheeky and not encouraging. Some users might find it offensive.	The two participants who used this application were indifferent to the performance feedback.
	OpenSpell	
1.3	The use of the labels <Say>, <Guess>, and <Spell> are not descriptive of their functionalities.	The two participants who used this program selected the <spell> menu option when asked to learn the spelling of a few words. However, this functionality is provided within <say> menu option.
1.13, 3.7, 4.3	When the <Spell> option is selected, the user cannot use the keyboard to provide input but must use onscreen keyboard. There is no instruction stipulating this restriction.	When asked to do some spelling exercises, both participants first attempted to use the keyboard to provide their input, only to realise later that they can only use the onscreen keyboard.
1.3	The * symbols used to represent the level of difficulty are not intuitive	Only one of the two participants was able to associate the * symbols with the level of difficulty. The other user did not know how to set the difficulty level.
	Themba's Journey	
3.2, 4.9	The narration voice is only in Xhosa. Non-Xhosa users who cannot read cannot use the application.	None
3.9, 4.9	To access an English version, the user must hover the mouse on the speech bubble. This can be problematic for users with limited use of their hands. The information provided under <Help> did not specify this.	Both participants did not know how to get the English version of the application until they were told. Much effort was required by participants to move pointer around the speech bubbles in order to read English versions.

3.6	None of the expert evaluator identified the dark background of Themba's Journey as potential problem. This is due to the close-up environment where the evaluation was conducted.	Application background was very dark. Both participants had to shield their faces and screen with hands. The dark background was made worse because the DD is located in an open space with excessive natural lighting and glaring sun.

The two problem sets were also matched with the evaluation heuristics in Table 2.

The results presented in Table 3 illustrate the value of using a combination of heuristic evaluation and field usability evaluation methods. The two methods each yielded errors that were overlooked in the other method. The heuristic evaluation method highlights what experts believe could constitute potential problems for users while the field evaluation revealed actual problems that impacted on users' tasks. There was also high variation in the nature of the usability/accessibility problems recognized by evaluators, with 48% of problems reported by single evaluators. The differences in the type of problems identified by evaluators has been termed the evaluator effect by authors like Hornbaek and Frojaer [16] and Hertzum and Jacobsen [15].

We also found that the heuristic evaluation method in general, yielded the kind of usability errors that can be seen as predictable; evaluation with real users produced additional errors that are generally unpredictable. An example of this occurred in the use of the terms 'username' and 'surname', where the experts could distinguish between the two, whilst some users confused the two.

The heuristic evaluation method identified a large number of usability and/or accessibility problems (71), some of which were low-severity problems that may not necessarily affect users' tasks. Examples of these includes the lack of consistency in the use of upper and lower case letters for interface elements captions and the layout of log-in instructions. Eight additional problems which surfaced in the field study were not recognized by expert evaluators. For example, the low support for users to identify and locate the fields in the registration form that caused error messages. This resulted in an unintended deletion of input data.

Although some of the concerns raised by expert evaluators were inconsequential for the participants in this study (e.g. the small instruction font size and the sarcastic performance feedback), they could constitute problems for other users. Attention to as many of the potential problems as possible will improve general usability for a wider spectrum of users.

The use of questionnaires as a secondary instrument during the field study to triangulate data showed the kind of inconsistencies that may arise between what users say and what they actually do. For example, responding that they understood the meaning of symbols and icons when in reality they were unable to determine their purpose.

One of the shortcomings of the heuristic evaluation method is that of overlooking real user problems. A field usability evaluation on the other hand may not necessarily reveal all the problems that potential users may encounter due to the relatively small number of participants and their level of expertise in the particular study. By combining the heuristic evaluation method with a field usability evaluation, we were able to address these limitations. Furthermore, the kind of disruptions that is characteristic of field evaluation settings was minimized by scheduling the evaluations after the official closing hour of the school.

Overall, both the heuristic evaluation method, using a set of heuristics specifically derived to suit such environment, and the field usability evaluation methods were found to be appropriate methods for assessing the usability of a non-standard interactive system.

6. CONCLUSION

In this paper we presented the use of multi-category heuristics in the evaluation of a non-standard system, the DD. We described how the heuristic evaluation method was complemented with field usability evaluation at a local school where the real context of the system's use is retained. We showed that despite the limitations inherent in different evaluation methods, using a combination of evaluation methods can offset such limitations.

Future research will involve field evaluation with a larger and more diverse user group to assess the extent to which the problems that were identified by expert evaluators, and which were not experienced by participants in this study, are reflected in the use of the DD by other user groups.

7. ACKNOWLEDGMENT

We thank the DD team at Meraka Institute for allowing us to evaluate the system. We also wish to express our gratitude to the expert evaluators for making time out of their busy schedule to evaluate the DD, as well as the learners who participated in the field usability study.

8. REFERENCES

[1] Adebesin, T. F. 2010. Report on the Usability and Accessibility Evaluation of the Digital Doorway. Available from http://hufee.meraka.org.za/Hufeesite/links/files/Technical%20Report%20-%20Digital%20Doorway%20Evaluation-Protected.pdf/view

[2] Alessi, S. M., & Trollip, S. R. 2001. Multimedia for Learning: Methods and Development (3rd ed.). Massachusetts: Allyn & Bacon.

[3] Barnum, C. M. 2002. Usability Testing and Research. Massachusetts: Allyn & Bacon.

[4] Cockton, G., Woolrych, A., & Lavery, D. 2008. Inspection-Based Evaluations. In A. Sears & J. A. Jacko (Eds.), *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications* (2nd ed.). New York: Lawrence Erlbaum Associates.

[5] Davis, F. D. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3), 319-340.

[6] Dix, A., Finlay, J., Abowd, G. D., & Beale, R. 2004. Human-Computer Interaction (3rd ed.): Pearson Education Ltd.

[7] Dumas, J. S. 2003. User-Based Evaluation. In J. A. Jacko & A. Sears (Eds.), *The Human-Computer Interaction Handbook*. Mahwah: Lawrence Erlbaum Associates.

[8] Gardner-Bonneau, D. 2010. Is Technology Becoming More Usable -or Less - and With What Consequences. Journal of Usability Studies, 5(2), 46-49.

[9] Gelderblom, J. H. 2008. Designing Technology for Young Children: Guidelines Grounded in a Literature Investigation on Child Development and Children's Technology. PhD Thesis, UNISA. Retrieved from http://etd.unisa.ac.za/ETD-db/theses/available/etd-09172008-132111/unrestricted/thesis.pdf

[10] Gray, W. D., & Salzman, M. C. 1998. Damaged Merchandise? A Review of Experiments That Compare Usability Evaluation Methods. Human-Computer Interaction, 13, 203-261.

[11] Gush, K., Cambridge, G., & Smith, R. 2004. The Digital Doorway - minimally invasive education in Africa. Paper presented at the Proceedings of the ICT in Education Conference, Cape Town.

[12] Gush, K., De Villiers, M. R., Smith, R., & Cambridge, G. In Press. Digital Doorways. In J. Steyn, J. Belle & E. M. Villeneuva (Eds.), *Development Informatics and Regional Information Technologies: Theory, Practice and the Digital Divide* (Vol. 2). Pennsylvania: IGI.

[13] Henry, S. L. 2002. Understanding Web Accessibility. In J. Thatcher, P. Bohman, M. R. Burks, S. L. Henry, B. Regan, S. Swierenga, M. D. Urban & C. D. Waddell (Eds.), *Constructing Accessibility Web Sites*. Birmingham: Glasshaus.

[14] Henry, S. L. 2007. Just Ask: Integrating Accessibility Throughout Design. Retrieved from http://www.uiaccess.com/JustAsk/

[15] Hertzum, M., & Jacobsen, N. E. 2003. The Evaluator Effect: A Chilling Fact About Usability Evaluation Methods. International Journal of Human-Computer Interaction, 15(1), 183-204.

[16] Hornbaek, K., & Frokjaer, E. 2008. A Study of the Evaluator Effect in Usability Testing. Human Computer Interaction, 23(3), 251-277.

[17] IBM. 2009. IBM Software Accessibility Checklist Retrieved 24 November 2009, from http://www-03.ibm.com/able/guidelines/software/accesssoftware.html

[18] International Organization for Standardization. 1998. ISO 9241-11 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 11: Guidance on Usability.

[19] Iwarsson, S., & Stahl, A. 2003. Accessibility, Usability and Universal Design - Positioning and Definition of Concepts Describing Person-Environment Relationships. Disability & Rehabilitation, 25(2), 57-66.

[20] Jeffries, R., Miller, J. R., Wharton, C., & Uyeda, K. 1991. User Interface Evaluation in the Real World: A Comparison of Four Techniques. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

[21] Lindgaard, G. 1994. Usability Testing and System Evaluation: A Guide for Designing Useful Computer Systems. London: Chapman & Hall Computing.

[22] Malone, T. W. 1980. What Makes Things Fun to Learn? Heuristics for Designing Instructional Computer Games. Paper presented at the 3rd ACM SIGSMALL Symposium and 1st SIGPC Symposium on Small Systems.

[23] Malone, T. W. 1981. Toward a Theory of Intrinsically Motivating Instruction. Cognitive Science, 5(4), 333-369.

[24] Mayhew, D. J. 1992. Principles and Guidelines in Software User Interface Design. New Jersey: Prentice Hall, Inc.

[25] Molich, R., & Dumas, J. S. 2008. Comparative Usability Evaluation (CUE-4). Behaviour & Information Technology, 27(3), 263-281.

[26] Nielsen, J. 1993. Usability Engineering. Boston: Academic Press, Inc.

[27] Nielsen, J. 1994. Heuristic Evaluation. In J. Nielsen & R. L. Mack (Eds.), *Usability Inspection Methods*. New York: John Wiley & Sons, Inc.

[28] Nielsen, J. 2003. Usability 101: Introduction to Usability Retrieved 20 July 2009, from http://www.useit.com/alertbox/20030825.html

[29] Norman, D. A. 2001. The Design of Everyday Things. London: MIT Press.

[30] Preece, J., Rogers, Y., & Sharp, H. 2007. Interaction Design: Beyond Human-Computer Interaction (2nd ed.). Chichester: John Wiley & Sons Ltd.

[31] Pühretmair, F., & Miesenberger. 2005. Making sense of Accessibility in IT Design – Usable Accessibility vs. Accessible Usability. Paper presented at the Database and Expert Systems Applications (DEXA’05).

[32] Rogoff, R. 2001, 24-27 Oct 2001. Making Electronic Information Accessible to Everyone. Paper presented at the Professional Communication Conference, (IPCC 2001).

[33] Rubin, J. 1994. Handbook of Usability Testing: How to plan, design, and conduct effective tests. New York: John Wiley & Sons.

[34] Sears, A. 1997. Heuristic Walkthroughs: Finding the Problems Without the Noise. International Journal of Human-Computer Interaction, 9(3), 213-234.

[35] Shelley, B. 2001. Guidelines for Developing Successful Games Retrieved 10 March 2010, from http://jnoodle.com/careertech/files/GuidelinesDevelopingSuccessfulGames.pdf

[36] Shneiderman, B. 1998. Designing the User Interface: strategies for Effective Human-Computer Interaction (3rd ed.). New York: Addison Wesley.

[37] Story, M. F., Mueller, J. L., & Mace, R. L. 1998. The Universal Design File: Designing for People of All Ages and Abilities Retrieved 17 November 2009, from http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/19/ac/11.pdf

[38] United States Access Board. 2000. Electronic and Information Technology Accessibility Standards (Section 508) Retrieved 09 November 2009, from http://www.access-board.gov/sec508/standards.htm

[39] Vanderheiden, G. C. 1994. Application Software Design Guidelines: Increasing the Accessibility of Application Software to People with Disabilities and Older Users Retrieved 04 November 2009, from http://trace.wisc.edu/docs/software_guidelines/software.htm

[40] Wharton, C., Braffort, J., Jeffries, R., & Franzke, M. 1992. Applying Cognitive Walkthroughs to More Complex User Interfaces: Experiences, Issues, and Recommendations. Paper presented at the SIGCHI Conference on Human factors in Computing Systems.

[41] World Wide Web Consortium. 1999. Web Content Accessibility Guidelines 1.0 Retrieved 7 August 2009, 2009, from http://www.w3.org/TR/WAI-WEBCONTENT/

The Impact of Usability on Efforts to Bridge the Digital Divide

Funmi Adebesin
Meraka Institute of the CSIR and School of Computing, UNISA
P.O. Box 395
Pretoria, 0001, South Africa
fadebesin@csir.co.za

Paula Kotzé
Meraka Institute of the CSIR and School of Computing, UNISA
P.O. Box 395
Pretoria, 0001, South Africa
paula.kotze@meraka.org.za

Helene Gelderblom
School of Computing, UNISA
P.O. Box 392
UNISA, 0003, South Africa
geldejh@unisa.ac.za

ABSTRACT

There is growing efforts to narrow the digital divide both locally and internationally. One such effort is the Digital Doorway project driven by the Department of Science and Technology (DST) and the Meraka Institute of Council for Science and Industrial Research (CSIR). It involves a non-standard computer system housed in a rugged, custom-designed kiosk. The preloaded software applications run on the Ubuntu Linux operating system, but the interface is not standard Linux. The project has mainly focused on providing physical access to computers in underprivileged communities around South Africa, without any formal usability evaluation of the software installed on the system. Our belief is that unless basic usability concerns are addressed in these types of development projects, the dream of the providing effective access may remain just that – a dream. This paper highlights the important role that usability plays in the drive towards narrowing the digital divide. We report on the outcome of a usability evaluation field study conducted on the Digital Doorway. The results suggest that there is a need for in-house usability standards to guide the various developers (in-house or external) who build applications for the Digital Doorway.

KEYWORDS

Digital divide, field evaluation, ICT for development, usability, usability evaluation.

1 INTRODUCTION

In the contemporary information revolution age, the use of information and communication technologies (ICTs) continues to influence every aspect of our lives. With the explosive growth of the World Wide Web and the Internet, the web is becoming an essential portal to access and share information and conduct business transactions. However, many are being excluded from the potential economic and social benefits of new technologies.

To address the problem, international and national initiatives are ongoing to provide access to technologies with the aim of bridging the digital divide. Many of the efforts to narrow the divide have been concerned with the provision of physical ICT devices. Examples of such projects includes the one laptop per child project [http://www.laptop.org/en/] and the Digital Doorway initiative by the

Department of Science and Technology (DST) and the Meraka Institute of Council for Science and Industrial Research (CSIR).The usability of these devices and applications installed on them, however, constitute one of the crucial factors to effectively narrow the divide [Nielsen, 2006].

One measure of the success of these initiatives, among other factors, is the ease of use of the computer devices by users [Davis, 1989; Nielsen, 2003]. It is therefore crucial that designers incorporate usability design principles early on in the design process. When the target user group has special needs this becomes even more important.

Usability is generally defined in terms of an application's effectiveness, efficiency and the satisfaction of the user. Every interactive system should be evaluated to (i) determine the ease of use of the systems' functionalities (ii) assess the user interaction experience, and (iii) identify any specific problems in the system [Dix, Finlay, Abowd and Beale, 2004].

There are several factors contributing to digital divide, among them financial constraints, lack of adequate skills and complexities of the interfaces of ICT devices, i.e. their usability. Currently, research focusing on usability, as an area that can be exploited in the effort to narrow digital divide, is limited. The purpose of the paper is to address this gap.

The Digital Doorway, a non-standard[1] computer system was first deployed in the rural community of Cwili in 2002. Since then, the Digital Doorway project has mainly focused on providing physical computers to underprivileged communities around the country, without any formal usability evaluation of the software applications installed on the systems. We describe here a field usability evaluation conducted to determine how easy the Digital Doorway is to use by users with limited computer-literacy.

The rest of the paper is structured as follows: in section 2, we provide a formal definition of digital divide and describe the different aspects of the divide. Section 3 briefly introduces the concept of usability while section 4 examines previous studies that focused on the digital divide from a usability perspective. In section 5, we provide an overview of the Digital Doorway, the target system evaluated in this study. The discussion on how the Digital Doorway was evaluated to determine its usability, at a local school, and the results obtained from the evaluation is provided in section 6. We discuss the role of usability in the efforts to narrow digital divide in section 7 and conclude the paper in section 8.

2 THE DIGITAL DIVIDE

Digital divide is a multidimensional phenomenon that refers to the disparity in access, distribution, and use of ICTs between two or more populations [Wilson, 2006]. It affects different age and gender groups, communities, races and regions of the world [Camacho, 2005]. The divide can also be seen

[1] Non-standard in this context means systems that do not display standard operating system interfaces or use standard equipments.

among different population groups within the same nation. For example, in the United States, white and Asian people are over 20% more likely to own computers than their black and Hispanic counterparts [Cooper and Kugler, 2009]. Closer to home, in South Africa, only 2% of black households had computers in 2001, compared to 46% of white households [Statistics South Africa, 2001][2]. This can be attributed to the legacy of apartheid and economic exclusion which have resulted in huge disparity between the black and white population groups [Martindale, 2002]. The 2007 community survey, conducted by Statistics South Africa, showed general increase in the ownership of household computers from 8.6% of the population in 2001 to 15.7% in 2007 [Statistics South Africa, 2007]. However, the report did not provide a breakdown of household computer ownership among the various population groups.

Digital divide is not only about the acquisition of computing devices. Other factors that contribute to the widening of the divide includes [Wilson, 2006]:

- *Financial constraints*: This refers to the inability of individuals, communities or governments to acquire ICT devices and sustain payments to service providers. For poor communities, where the primary concern is the ability to feed their families, ICT devices cannot be afforded.
- *Lack of adequate cognitive resources*: Effective interactions requires the user to possess the basic ICT skills that will enable him/her to recognize the need for information, find the information, process and evaluate the information for its appropriateness, and utilize it in a meaningful way.
- *Complex interface designs (usability)*: Even when ICT devices are available, the complexity of the interfaces makes it impossible for the novice user to access the content. Other aspects of usability involve accessibility to people with special needs, such as the disabled and the elderly.
- *Lack of relevant content*: Another factor contributing to digital divide is the lack of content that are locally and culturally relevant. The predominant language of the Internet for example, is still English. According to the 2009 estimate by the Internet World Statistics [2009], English language ranked highest among the top ten Internet users by language, with no African language featuring among them (see Figure 1). From the perspective of a user in a developing country, content access in the local language is one of the critical requirements for bridging digital divide.

[2] The latest official numbers issued by Statistics South Africa.

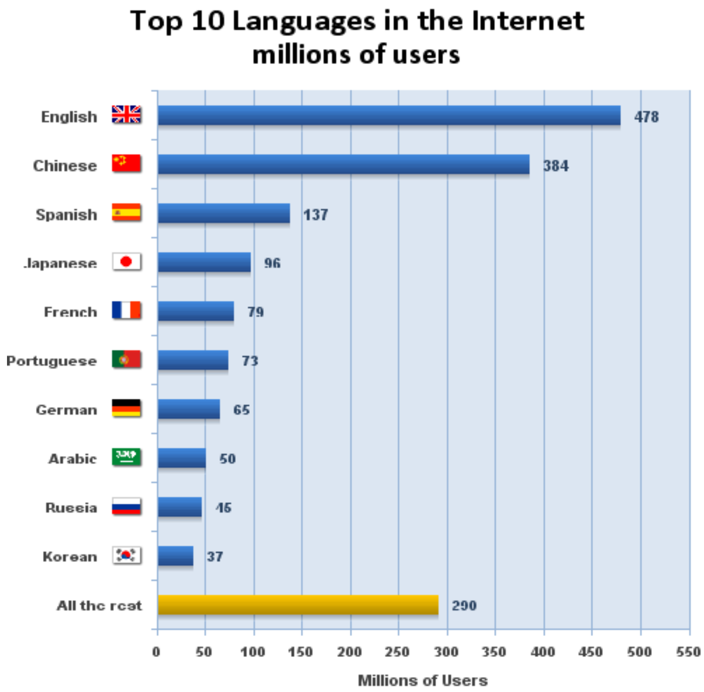


Figure 1: Top Ten Internet Languages [Internet World Stats, 2009]

3 USABILITY

Usability is defined by the International Organization Standardization (ISO) [1998] 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".

Usability is one of the focus areas of human-computer interaction (HCI), a field of study concerned with the design, implementation, and evaluation of interactive systems taking into account the context of use and the task the user needs to accomplish. Usability cannot be retrofitted into a design later in the development life cycle. Usability specifications should form part of the requirements specification process [Dix et al., 2004].

One method for incorporating usability into the design process involves the use of usability principles. These principles can guide designers so that their design decisions do not negatively affect the usability of the application. Usability design principles can be widely applied to a variety of situations as they impose fewer constraints in terms of how the principles should be implemented [Dix et al., 2004; Kotze and Johnson, 2004]. For example, the principle of feedback states that adequate feedback should be provided to users to enable them determine what they need to do next in order to complete the task at hand. However, the principle is flexible about how to provide the feedback [Kotze and Johnson, 2004; Preece, Rogers and Sharp, 2007]. For instance, feedback could be provided using text, graphics, or audio output or a combination of these, depending on the requirements of the specific user groups.

4 DIGITAL DIVIDE AND USABILITY

A search for research articles using the keywords ‘digital divide and usability’ and ‘ICT for development and usability’ do not yield a large harvest. Our search produced only a few studies that reported on usability evaluation of ICT devices and applications deployed specifically with the aim of narrowing digital divide. Below we review those whose findings are most relevant to our study.

Researchers such as Fuchs and Horak [2008] and Gebremichael and Jackson [2006] merely mentioned usability as one of the factors contributing to digital divide. The report by Boeltzig and Pilling [2007] addressed several factors (including usability) that impacted on the ability of specific user groups, such as the elderly and the disabled, to access and make effective use of electronic government services. Shneiderman [2001] provided an overview of the first ACM conference on universal usability held in November 2000, where participants identified universal usability as one of the strategies to narrow the widening digital divide.

brigdes.org [http://www.bridges.org/Real_Access] identified twelve evaluation criteria for determining why development projects aimed at narrowing digital divide sometimes fail to achieve their goals. Among these criteria is the appropriateness of the technology for the intended local community. The measure of technology appropriateness includes energy requirements, security of the devices and the ease of use, i.e. usability, of the interface between the user and the devices.

Liu and Meng [2007] conducted a study on the usability of mobile phone among off-farm workers in China. Off-farm workers are people who leave their farms in the rural areas in search of other forms of employment opportunities in the cities [Nielsen, Smyth and Zhai, 2010].

The authors found that while the study participants were eager to embrace new technologies to improve their conditions, lack of considerations for the special usability requirements of low-literacy users by designers prevented them from taking advantage of the opportunities offered by new technologies. Although over 90% of off-farm workers possessed mobile phones, the majority of the study participants were merely using their mobile phones to make and receive telephone calls. Other useful functionalities, such as the phonebook feature, were never used. Rather than use this feature to store the details of potential employers for example, they wrote these down in pieces of paper, with the risk of misplacement.

A set of representative tasks were given to the participants, for example, changing a phone’s ringtone and retrieving previously stored phone numbers. Results from the study showed that participants were unable to interpret the meaning and functionality of the features required to complete the tasks. Majority of the participants required assistance from the evaluator; they made large number of mistakes and spent considerable amount of time to complete the tasks.

In a survey on the use of electronic information systems among low income and underserved Americans, Lazarus and Mora [2000] reported that lack of locally relevant content and usable interfaces formed part of the barriers to these user groups' taking advantage of the opportunities offered by new technologies.

The studies above all referred to usability as part of the factors that could help narrow digital divide, but few studies have been done to formally evaluate the usability of the applications developed specifically for this purpose. Only the study by Liu and Meng [2007] specifically focused on evaluating the usability. Our aim with this paper is to raise awareness and stir the debate on usability and the role it can play in narrowing the digital divide.

5 OVERVIEW OF THE DIGITAL DOORWAY

This section briefly introduces the Digital Doorway, the target system evaluated in a field usability study. First we provide the background to the development and motivations behind the Digital Doorway project and then describe the features and functionalities of the interfaces and applications selected for evaluation.

The Digital Doorway project is a joint initiative by the DST and the Meraka Institute of CSIR. Digital Doorways are non-standard computer systems housed in rugged, custom-designed kiosks with multiple terminals that can be accessed simultaneously by users. The terminals are equipped with metal keyboards and reinforced touchpad for user input. The robust housing and metal keyboard is necessitated by the need to protect the system against acts of vandalism. The applications and content, which run on the Ubuntu Linux operating system, are preloaded [Gush et al., In Press]. However, the interface does not follow any particular design standard or operating system interface.

The Digital Doorway project is based on the 'hole in the wall' concept from India [Mitra and Rana, 2001], and aims to promote computer literacy through unassisted learning. It is an attempt to narrow the digital divide [Cambridge, 2008; Gush et al., In Press] by installing the computers in underprivileged communities such as schools, police stations and community centres around South Africa. Till date, 206 Digital Doorways have been deployed around the country.

The Digital Doorway provides extensive access to software applications and other resources, the majority of which are open source or third-party applications. These includes the OpenOffice suites, educational games, scientific simulations, Wikipedia documents and Mindset applications – a South African curriculum-based educational program [Gush, Cambridge and Smith, 2004].

Applications developed in-house are sometimes implemented by contract software developers. Currently, the Digital Doorway does not provide support for the use of assistive devices, such as screen readers for visually impaired users. A three-terminal Digital Doorway is shown in Figure 2.

Figure 2: A three-terminal Digital Doorway

This evaluation study focused on the usability of the interfaces and applications developed in-house specifically for the Digital Doorway. The specific interfaces and applications evaluated are: the login screen, the new user registration form, the main desktop, and three educational games - *What-What Mzansi*, *OpenSpell*, and *Themba's Journey*. The following subsections provide brief descriptions of the interfaces and applications evaluated.

5.1 The Digital Doorway Login Screen

The login screen is the first interface between the user and the Digital Doorway. Users access content by logging in as a guest user, a registered user, or by creating a new user account and then logging in using the newly created account. The main language on the login screen is English, but equivalent information is available in four other South African languages, namely Xhosa, Afrikaans, Sotho and Venda. A guest user can simply access content of the Digital Doorway by typing 'dd1' in the username textbox. A new user account is created by typing 'new' in the username textbox; this will activate the registration form.

5.2 The New Account Registration Form

Users may chose to create new user accounts by completing a simple electronic form. Items on the form are organized into two main groups – 'Personal Details' and 'User Details'. Within the personal details group, demographic information such as name, age, and gender are provided. User-selected username and password are chosen within the user details group. The form also provides users with hints on the type of data expected at certain fields, for example the password field. The form requires all data fields to be filled, although this is not explicitly specified in the form. After completing the form, the information provided is stored by clicking on the <Register User> button.

5.3 Digital Doorway Desktop

Following a successful login, applications and content of the Digital Doorway can be accessed by clicking on icons on the desktop or by selecting from the two menu options 'Programs' and 'Resources'. The desktop also provides global volume control either by clicking on a 'volume control' icon or through a more advanced volume control window. Users can log out of the system by clicking on an ‘exit’ button (designated by a right pointing arrow ⇒) or from the ‘System’ menu.

5.4 What-What Mzansi

What-What Mzansi is an educational quiz game in the form of yes/no questions. Developed to provide content relevant to the South African environment, the program provides two levels of difficulty – <Easy> and <Advanced>. Context-specific instructions are provided when the user clicks on the <?> icon, located at the top right corner of the screen while the <X> icon closes the application. The interface provides three menu options. <About> menu presents the user with information on the Digital Doorway project and its achievements, together with details of the game developers. The questions are asked and answered when <play> is activated, while <hi-Scores> lists the scores of the top ten registered users. On the selection of a difficulty level, a local voice welcomes the player and reads out the questions which can be answered by clicking on <Yes> or <No>. Each session lasts 60 seconds. The score for each question can range from 2 to 10, depending on how fast the player answers it. The interface of What-What Mzansi is shown in Figure 3.

Figure 3: Interface of What-What Mzansi

5.5 OpenSpell

OpenSpell is an educational spelling game that is available in all eleven South African official languages. It provides three levels of difficulty designated with *, **, ***. The interface, shown in Figure 4, includes an onscreen keyboard used for providing input in spelling exercises. The program provides three menu options. Clicking on <say> brings up a series of pictures of words to be spelt. For each word, a voice in the chosen language speaks out each letter as well as its pronunciation. <Guess> is based on the hangman word guessing game, while the <spell> option tests the users’ spelling skills. Spelling exercises is done by clicking letters from onscreen keyboard. Users are given two opportunities to spell words, after which the correct answer is provided.

Figure 4: OpenSpell interface (with the <say> menu active)

5.6 Themba's Journey

Themba's Journey is a life skills program in the form of a narrative. Themba is a young man who makes a journey from his village to the city of Johannesburg in search of a job. Throughout the narration, Themba is faced with potential life-changing situations that require him to make choices. The user, who assumes the role of Themba, must make these choices for him, for example, whether or not to take drugs. Each decision can result in positive or negative consequences. The interface (Figure 5) provides three menu items. The <Help> menu contains the navigation and game instructions. The main story is narrated within the <Play> environment. Clicking on <Exit> closes the program.

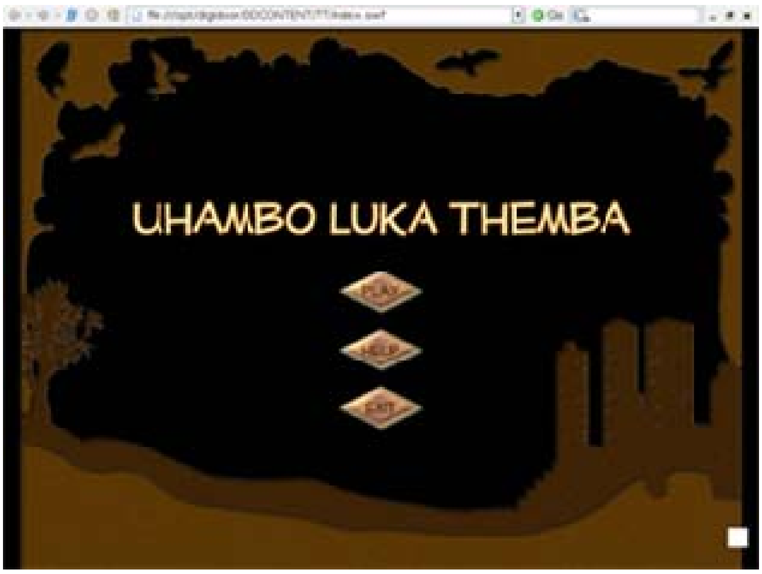


Figure 5: Themba's Journey Interface

6 A FIELD USABILITY EVALUATION OF THE DIGITAL DOORWAY

Designers typically make assumptions about users [Gardner-Bonneau, 2010] and usually develop applications for the so-called average user. The reality is that users are quite diverse in terms of their age, gender, expertise, ability, and nationality [Kotze and Johnson, 2004; Norman, 2001; Shneiderman, 2000]. The validity of designers' assumptions needs to be tested by evaluating with real users.

This section discusses the field usability evaluation conducted at a local school where the Digital Doorway is installed, as well as the results obtained from the evaluation.

6.1 The Evaluation Environment

We evaluated the Digital Doorway using the field observation method at a local secondary school where the context of the system's use is retained. Our choice of a school as an evaluation venue among other potential centres (e.g. community centre and police stations) having Digital Doorway installed was based on two factors (i) the three applications evaluated were educational games, hence it makes sense to evaluate the usability of the applications among school children (ii) a study on the usage patterns of the Digital Doorway at a number of representative centres around South Africa recorded secondary schools as having the most successful usage [Gush and De Villiers, 2010].

At this particular center, the Digital Doorway is installed in an open area on one of the school's corridors to provide unrestricted access to users. Children from surrounding homes also have access to the Digital Doorway soon after the school closes until 18:00 late in the afternoons. Although the provision of unrestricted access to the system is commendable, there is inadequate provision of shading from sun glare.

Prior to conducting the study, formal approval was obtained from the school principal. Parents/guardians of participants also signed informed consent forms.

Nine learners participated in the evaluation, six of whom were given pre-defined tasks to complete while the other three participants were allowed to use the system as they wished. Of the six participants given pre-defined task, two participants each used one of the educational games. These participants were also required to register a new user account before accessing the applications unless they had a valid account. The other three participants could access the system as a guest if they wished. The profile of the participants, together with the applications they used is provided in Table 1.

Table 1: Profile of Field Evaluation Participants

Participants	Age	Gender	Application used
1	17	F	OpenSpell
2	13	M	OpenSpell
3	16	M	What-What Mzansi
4	15	M	What-What Mzansi
5	15	M	(Free Exploration) What-What Mzansi and Four-in-a-row game
6	14	M	Themba's Journey
7	18	F	Themba's Journey
8	15	F	(Free Exploration) Themba's Journey
9	13	M	(Free Exploration) KTuberling and Penguin games

6.2 Evaluation Process

In a conventional field study, participants are observed as they carry out normal or routine activities using the target system in the natural context of use either at home or the workplace. The natural context allows the observer to see the actual ways in which the system is being used; thus revealing some details that may be difficult to obtain if another evaluation method, such as the heuristic evaluation method, was used [Dumas, 2003].

In this study, we modified the field evaluation by giving some of the participants pre-defined tasks to complete and allowing other participants to freely explore the system. This enabled us to focus the evaluation on the specific interfaces and applications identified for the evaluation while at the same time allowing us to observe the type of applications the learners typically access. A sample pre-defined task list is shown in Figure 6.

To avoid disruptions to learning activities as much as possible, and minimize distractions from noise, evaluation sessions took place in the afternoons well after the official closing hour of the school.

Using the cooperative evaluation style, participants were encouraged to ask questions and assistance whenever they got stuck with any activity. The field evaluation facilitator, for example, provided subtle hints and assistance after allowing participants sufficient time to attempt to locate an interface element without success. This approach is justified since the Digital Doorway is not a transaction processing system where the speed of task completion is a measure of usability.

Nine evaluation sessions were conducted over a two-week period. Each session lasted between thirty and forty-five minutes. The sessions were recorded on video cameras after assurance to participants of their anonymity. The facilitator also took notes of important events as they occurred.

After each session, footage of the evaluation was reviewed and compared with the facilitator's note in order to check for any inconsistency between the two, before preparing for the next day's evaluation session. This was to ensure that data from the sessions were not mixed-up.

Digital Doorway Evaluation: Task list

1. Read the screen instruction on how to register as a new Digital Doorway user if you are not a registered user.
2. Complete the registration form if you are not a registered user, otherwise proceed to step 3.
3. Start the Digital Doorway by providing the requested information.
4. Search for the spelling game '**OpenSpell**'.
5. Remember to provide verbal feedback all the time.
6. Search for and read the instruction on how to play the game.
7. Choose how challenging (difficult) you want the game to be.
8. Learn how to spell a few words.
9. Change the volume to suit your need.
10. Do some spelling exercises.
11. Do a few guessing exercises.
12. Change the language to another one of your choice.
13. Close the Digital Doorway when you are done.

Figure 6. Task list for field usability evaluation (using the application 'OpenSpell')

6.3 Evaluation Results

Various usability problems were encountered by users and observed by the field evaluation facilitator. All of these were software usability problems except one which was a hardware problem. Some of the problems affected the completion of participants' tasks while others constituted a source of minor irritations to them.

The total number of usability problems found during the field evaluation was thirty-eight. Thirty-four of these were software problems that affected task execution by participants; three were a source of irritation to the participants while two were hardware problems. Analysis of the specific interfaces and application in which the problems were located (Figure 7) showed that six problems were related to the login screen, eight affected the new user account registration form, while six involved the main desktop. In the educational game applications, four problems related to the quiz game, What-What Mzansi, six problems involved OpenSpell while a further six related to Themba's Journey.

As shown in Figure 7, a total of twenty problems were located in the login screen, the new account registration form and the main desktop. This constitutes 56% of the total software usability problems. At least two of these interfaces (the login screen and the main desktop) represent the first areas of contact between the user and the Digital Doorway. This is a concern for a system that aims to promote computer literacy through unassisted learning [Cambridge, 2008]. Successful and meaningful interaction begins with simple, easy to use and intuitive interfaces. We provide the description of the nature of problems revealed by the field usability evaluation in Table 2.

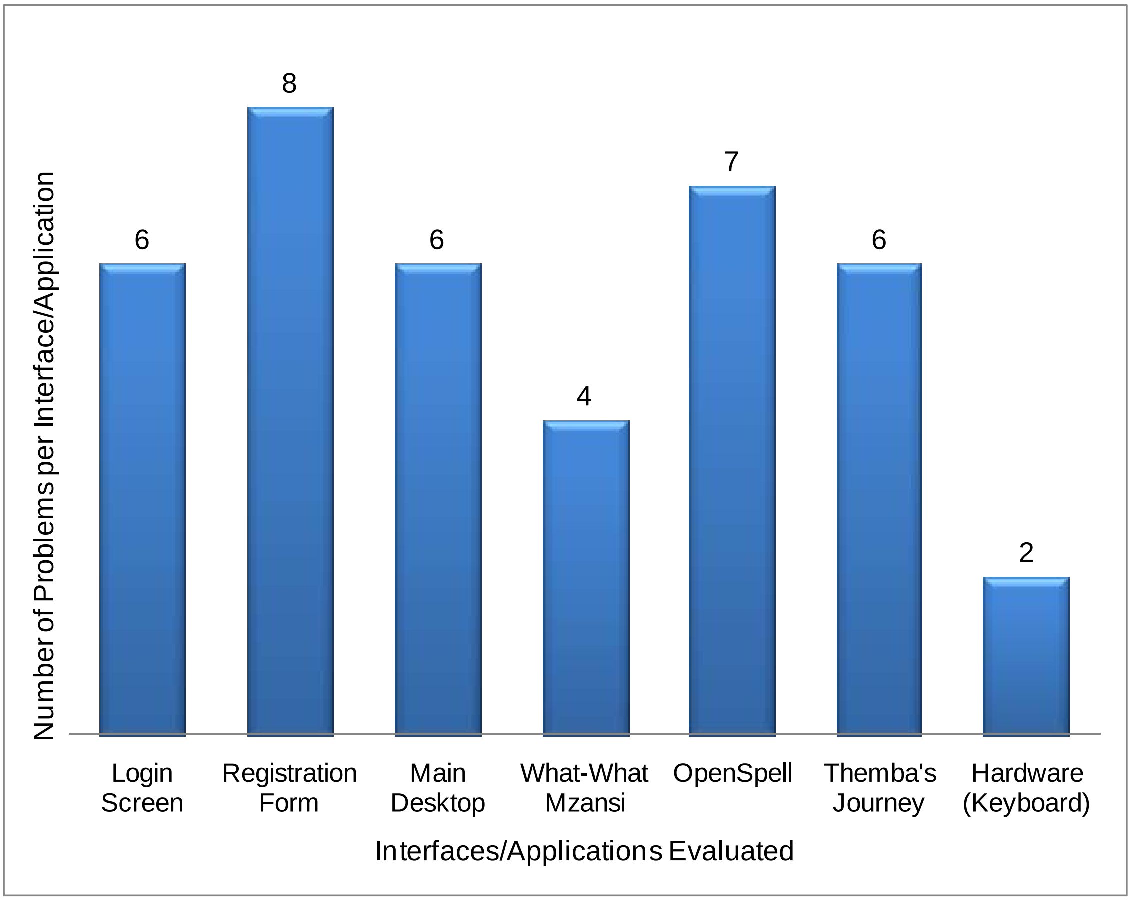


Figure 7. Number of Problems per Interface/Application

Table 2: Nature of Usability Problems

No	Problem Description
Login Screen	
1	Incorrect username and/or password were a common user error. The system returned the same screen over and over without an indication of what the mistake was. Most of the time the field observation facilitator had to inform the participants that the problem was with the username or password they were providing after a number of unsuccessful attempts to log in.
2	The system did not provide login information for users who had just registered or those with existing accounts. The only prompt on the screen is addressed to guest users and those wanting to create new user accounts. Some of the participants typed in 'new' or 'dd1' to log in, while others asked what should be done next before being told that the newly chosen username and password should be used to log in.
3	Participants sometimes confused their surname with a 'username' and typed their surname in the username field, though this was not the chosen username.
4	Some of the participants confused the <Enter> key on the keyboard with the key designated for producing a 'mouse click' effect because the keys were not labelled. However, after pressing one key without the desired effect they then pressed the other.
5	A screen resolution dialog box occasionally appeared and disappeared after a few seconds with the message 'For best picture quality change the resolution to 1024X. 1: Exit 2: Delete.' This made some of participants irritated as they did not know how to handle the information.
6	On a number of occasions, when participants were about to place the insertion point in the Username textbox, the following message appeared on a rollover 'Answer questions here and press Enter when done. For a menu press F10'. While this message did not seem to bother some of the participants in the

	study, its relevance is questionable.
New User Account Registration Form	
7	Some participants began typing their names only to realize later that the input was not being accepted and needed to place the insertion point within the first field before typing again.
8	Some of the participants input their name and surname in the 'Name and Surname' field without space in between the two. This common error will then bring up the following error message "Your name seems to be incomplete". Participants then spent some time trying to figure out what the problem is, sometimes without success until they were told what the problem is.
9	Participants typically kept the 'home language' and 'preferred language' fields empty only to have error messages urging them to fill the fields. None of the fields in the form is indicated as mandatory.
10	Some participants chose passwords with the length less than six characters. This resulted in the following error message "Passwords must be between 6 and 14 characters". This contradicted the hint provided next to the password field "6 to 12 numbers and letters".
11	While setting the password, a participant received the following error message 'The password contains illegal characters'. This participant could not comprehend the meaning of the error message. She had to ask the field observation facilitator for help.
12	The form did not facilitate the location of an error field. A participant erased his input in the password field accidentally, while trying to correct the name field entry following an error message. The insertion point remained in the password field after clicking on the <Register user> button. Without the participant realising this, he pressed the backspace key on the keyboard (←) several times and erased the wrong field unintentionally.
13	Two participants accidentally clicked on the <Cancel> button while intending to click the <Register User> button. This inadvertent user error resulted in the form being closed without any warning to the participant thereby erasing all the data fields input thus far. The toe buttons <Register User> and <Cancel> are located closely to each other on the form.
14	Three participants were unable to delete the wrong input in form fields until they were told how to. This task can only be accomplished by pressing a left pointing arrow key ←, which is not labelled, on the keyboard. This is actually a hardware usability problem that affected the use of the electronic form.
Main Desktop	
15	Only two of the six participants with pre-defined tasks found the location of the game applications on their own. Other participants unsuccessfully searched for the applications within the <Game> submenu, located in the <Resource> menu, before they were told where to find them.
16	Only three of the six participants with pre-defined tasks were able to locate the volume control buttons on the desktop, the other three required assistance after several failed attempts.
17	Four participants found the background colour to be too dark. On several occasions, they had to shield their faces and the screen with their hands while using the Digital Doorway to overcome the extent of reflection of the sun on the dark background.
18	A participant accidentally clicked on the ⇒ button, used to exit the system, while trying to locate the volume control button and the system was shut down without any warning.
19	Only three of the six participants given pre-specified tasks were able to log out of the system on their own without requiring assistance. One participant discovered the ⇒ button accidentally following an attempt to increase the volume output. The other two participants specifically asked for help following failed efforts to exit the system on their own. Of the three participants that explored the system as they wished, two knew the location of the ⇒ button while the other participant asked for help after unsuccessful attempts to exit on her own.
20	After clicking on the required game application icon, the screen will flicker and return to the Digital Doorway home page. Participants needed to click the icon several times before the game application was opened. This was frustrating to participants.
What-What Mzansi	
21	The two participants given pre-defined tasks using this application could not find the game instructions as required in the specified task. Intuitively, the two participants clicked on <about> menu option to search for the game instructions without success. This is because this menu contains information on the

	application developers and Digital Doorway project history and achievements.
22	At the start of the application, some of the control buttons and the character that reads out instructions and questions were hidden from user’s view. A full screen mode is activated by clicking arbitrarily around the taskbar. None of the two participants who used this application for the pre-defined task knew how to get the full screen view of the game. One participant, who explored the Digital Doorway as wished, chose What-What Mzansi. This participant was able to change to a full screen view without requiring any help.
23	Context-specific instructions are provided when a user clicks on <?> icon. However, none of the three participants who used this application accessed the information. Non-utilization could be (i) because they did not understand the functionality of this icon and (ii) because they never had the opportunity to select the icon as the questions were read immediately after the welcoming words. The main priority of these participants was to listen to the question and answer them.
24	One of the terminals used for the evaluation sessions had unusually large icons. This resulted in non-visibility of a number of control buttons, in this instance a right pointing arrow ‘>’ used for forward progression. This made it impossible for the participants to repeat the level which they had just completed as required following poor performance.
OpenSpell	
25	In similar pattern to the participants who used What-What Mzansi, the two participants that were required to use this application could not find the game instructions. Both participants clicked on the <about> menu option to search for the game instructions, without success. This is because this menu contains information on the application developers and Digital Doorway project history and achievements.
26	The two participants who used this program selected the <spell> menu option when asked to learn the spelling of a few words. However, this functionality is provided within the <say> menu option.
27	Only one of the two participants was able to associate the * symbols with the level of difficulty. The other user did not know how to set the difficulty level.
28	The quality of the voice output was poor even when volume was at the highest. Participants frequently had to keep their ears close to the screen.
29	One of the terminals used for the evaluation sessions had unusually large icons. This resulted in the taskbar covering the control buttons <Repeat> <Erase> <Enter> almost completely. One of the participants who used this terminal had to ask what should be done to ‘enter’ her input for a spelling exercise. On two occasions, the participant needed to erase incorrect inputs but due to none visibility of these buttons, she clicked on the <Enter> button. This was taken by the system as an incorrect answer. She was then prompted to try one more time as the application interpreted this as an incorrect answer.
30	When asked to do some spelling exercises, both participants first attempted to use the keyboard to provide their input, only to realize later that they can only use the onscreen keyboard.
Themba’s Journey	
31	The default language for this application is Xhosa. To access an English version, the user must hover the mouse on speech bubbles. The three participants who used the application (two with pre-defined tasks and one as a free system explorer) did not know how to get the English version until they were told.
32	Too much effort was required by participants to move the pointer around the speech bubbles in order to read English versions.
33	Application background was very dark. Participants had to shield their faces and screen with hands. The dark background is made worse because the Digital Doorway is located in an open space with excessive natural lighting and glaring of the sun.
34	Navigation instructions were provided in the <Help> menu. Although the participants read the instructions at the start of the session, they had forgotten about the functionality of some of these buttons, in this instance the <Skip> button by the time they were actually needed.
35	At the second crossroad, which was having the options ‘Walk’ and ‘Take taxi’, the ‘Walk’ option could not be executed. A participant had to select the ‘Take taxi’ option against her wish.
36	The main exit button was non-functional. Participants had to close the application with the browser exit button i.e. the <X> button.

Keyboard	
37	The primary purpose of the study was to evaluate a selection of interfaces and applications, however, the evaluation revealed a number of keyboard keys that were not functioning. These were: letters ‘K’, ‘L’, ‘O’, and ‘P’. This affected the choice of passwords selected by some participants.

The field usability evaluation identified problems that impacted on the successful completion of user tasks. Many of these were flagged as potential usability problems during an expert heuristic evaluation of the Digital Doorway [Adebesin, Kotze and Gelderblom, 2010], for example, the use of unintuitive icons and symbols. Others were not recognized as problems by expert evaluators because they were a direct consequence of the environment of use. An example of this kind of problem was the dark background, which was exacerbated by the reflection from the sun.

Other usability problems experienced by participants in the field usability evaluation revealed the Digital Doorway’s lack of error tolerance. This is a concern for a system that aims to promote computer literacy through unassisted learning and system exploration. For example two participants in this study unintentionally clicked on the <Cancel> button in the new account registration form. This button is located closely to the <Register User> button. Without any warning message, the form was closed, thus resulting in the loss of all the data provided by the participants. On another occasion, a participant accidentally clicked on the arrow button (⇒), used to shut down the system, while attempting to increase the volume for audio output. The system shut down without any warning, the participant had to log in again in order to complete the specified tasks.

7 USABILITY AS STRATEGY TO BRIDGE THE DIVIDE

The benefits of usability are enormous – easy to use interfaces, reduced error rate, less user frustration and ability to transfer knowledge from one application to similar ones, to name a few. Although concerns for interface usability are essential for all users, it is even more the case when the target user groups are inexperienced and underserved. These are the users groups where the gap of digital divide is widest. Inadequate design decisions by developers could negatively impact on these user groups’ ability to take advantage of the potential social and economic benefits of new technologies.

The results presented in Table 2 demonstrated the significant role that usability evaluation, especially with real users, can play in the drive to narrow digital divide. As discussed in section 2, digital divide is not only about the acquisition of ICT devices. Other aspects of the divide, which were revealed by the field usability study, are the following:

- *Lack of relevant content in the local language:* The availability of relevant content in the local language is one of the critical requirements to bridge digital divide [bridges.org, n.d; Wilson, 2006]. Although Themba’s Journey, one of the educational applications evaluated in this study, is provided in Xhosa and English languages, the usability of the English equivalent is affected by poor design decision that required users to hover the mouse pointer over speech bubbles to access the content. Another application, What-What Mzansi, is currently available only in the English language. Although the participants in this study were school children who understood English, the same cannot be said for other children in other rural South African locations.

- *Lack of cognitive resources and inadequate interface design:* In order to take advantage of potential benefits of ICT, the possession of basic ICT skills is essential. When users are lacking in these basic skills, effective interaction requires interfaces that are simple and intuitive.
  - The field usability evaluation showed the Digital Doorway's lack of concern for users in this regard. For example, only two of the six study participants given predefined tasks could locate the educational games What-What Mzansi, OpenSpell, and Themba's Journey. The first intuition of the participants was to check for these applications in the <Game> submenu, where even an experienced user would have expected to find them. However, these applications were placed inside a desktop folder named 'new_content'.
  - Other examples relating to inadequate interface design in the Digital Doorway is related to the use of symbols that do not adequately convey their functionality, even to the experienced users. (i) In OpenSpell, the educational spelling program, users are allowed to set the level of difficulty for the game. However, the method of implementation for this support in form of *s is flawed as designers cannot reasonably expect users to associate this symbols with the level of difficulty. Furthermore, nowhere in this application were users provided with instructions that could help them in determining the meaning of this symbol. (ii) On the desktop, the exit button is represented by a right pointing red arrow (⇒). The interpretation of the function of this symbol would have been difficult, even for the experienced user.

In a development environment such as the one in which the Digital Doorway project team operates, applications are typically implemented by contract and visiting developers. To overcome the problems identified during our usability evaluation, well-established usability guidelines could provide a solution. Although there are several well-established usability guidelines, for example the usability principles by Dix et al. [2004] and design principles by Norman [Norman, 2001], these basic principles have not been followed by developers due to lack of clear guiding principles and policies on usability and standardization. As an example of the standardization issue, in the education games What-What Mzansi and OpenSpell, users can exit the two applications by clicking the <X> button provided in a browser window, while the life skills program, Themba's Journey, provides an <Exit> button to close the application. Such inconsistencies will not allow users to transfer knowledge from one application to another.

Another reason could be that the developers are unaware of the existence of such usability guidelines or that they find them overwhelming. In such cases, a solution could be the establishment of in-house usability guidelines that are specific enough for the types of applications being implemented and which can provide guidance to these developers. There should also be processes in place to ensure that the guidelines are followed.

In addition to establishing in-house usability guidelines and standards, usability evaluation should be conducted with real users to assess the extent to which they can effectively use the system to accomplish their goals. As stated in section 1, there has been no usability evaluation of the software applications installed on the Digital Doorway over the past eight years. Efforts to bridge digital divide should not be concentrated only on providing physical access to technologies. Without proper usability, content that may potentially be of benefit may not be utilized.

When people do not possess the basic ICT skills to access the software, the interface should be particularly supportive and should facilitate learning by exploration. It should be tolerant of user error and designers should make every effort to hold the user's attention. An intuitive, easy to use interface will enable the underprivileged to take advantage of the economic and social benefits offered by new technologies.

8 CONCLUSION

In this paper we presented the results from a field study conducted to evaluate the usability of the Digital Doorway, a non-standard system deployed as part of the global efforts to narrow digital divide. We described the type of problems encountered by real users when using computer systems and demonstrated that lack of usability undermines the cost and effort to provide the underprivileged with technology. We cannot hope to narrow digital divide simply by making ICT devices available to disadvantaged people and not pay proper attention to the content. The usability of the interfaces of these devices is as important as the provision of the devices themselves.

In situations where applications are implemented by contract developers, as it is sometimes the case with applications installed on the Digital Doorway, the establishment of appropriate in-house usability guidelines will ensure that usability concerns are addressed by developers.

Our hope is that this paper will provide the impetus for people involved in projects aimed at narrowing digital divide to ensure that the devices are effectively utilized by the target user groups through the appropriate incorporation of basic usability principles in the design.

9 ACKNOWLEDGMENT

We thank the Digital Doorway team at Meraka Institute for allowing us to evaluate the system. We also wish to express our gratitude to the learners who participated in the field usability study.

10 REFERENCE

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)*

Boeltzig, H. & Pilling, D. (2007). *Bridging the Digital Divide for Hard-to-Reach Groups.* IBM Center for the Business of Government.

bridges.org (n.d). *Real Access / Real Impact Criteria*. Available from: http://www.bridges.org/Real_Access. Last access date: 31 May 2010

Camacho, K. (2005). Digital Divide. In Ambrosi, A., Peugeot, V. & Primenta, D. (Eds.) *Word Matters: Multicultural perspectives on information societies.* C & F Editions.

Cambridge, G. (2008). Digital Doorway: Enriching your Mind. *Proceedings of the 2nd CSIR Biennial Conference*

Cooper, J. & Kugler, M. B. (2009). The Digital Divide: The Role of Gender in Human-Computer Interaction. In Sears, A. & Jacko, J. A. (Eds.) *Human-Computer Interaction: Designing for Diverse Users and Domains.* Florida: Taylor & Francis Group, LLC.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly,* 13(3): 319-340.

Dix, A., Finlay, J., Abowd, G. D. & Beale, R. (2004). *Human-Computer Interaction*. Pearson Education Ltd.

Dumas, J. S. (2003). User-Based Evaluation. In Jacko, J. A. & Sears, A. (Eds.) *The Human-Computer Interaction Handbook.* Mahwah: Lawrence Erlbaum Associates.

Fuchs, C. & Horak, E. (2008). Africa and the Digital Divide. *Telematics and Informatics,* 25(2008): 99-116.

Gardner-Bonneau, D. (2010). Is Technology Becoming More Usable -or Less - and With What Consequences. *Journal of Usability Studies,* 5(2): 46-49.

Gebremichael, M. D. & Jackson, J. W. (2006). Bridging the Gap in Sub-Saharan Africa: A Holistic Look at Information Poverty and the Region's Digital Divide. *Government Information Quarterly,* 23(2006): 267-280.

Gush, K., Cambridge, G. & Smith, R. (2004). The Digital Doorway - minimally invasive education in Africa. *Proceedings of the ICT in Education Conference*

Gush, K. & De Villiers, M. R. (2010). Application Usage of Unsupervised Digital Doorway Computer Kiosks in Remote Locations in South Africa. *Proceedings of the Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT 2010)*

Gush, K., De Villiers, M. R., Smith, R. & Cambridge, G. (In Press). Digital Doorways. In Steyn, J., Belle, J. & Villeneuva, E. M. (Eds.) *Development Informatics and Regional Information Technologies: Theory, Practice and the Digital Divide.* Pennsylvania: IGI.

International Organization for Standardization (1998). *ISO 9241-11 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 11: Guidance on Usability.*

Internet World Stats (2009). *Internet World Users by Language*. Available from: http://www.internetworldstats.com/stats7.htm. Last access date: 06 April 2010

Kotze, P. & Johnson, C. (2004). *Human Computer Interaction Study Guide for INF120-8*. Pretoria: University of South Africa.

Lazarus, W. & Mora, F. (2000). *Online Content for Low-Income and Underserved Americans: The Digital Divide's New Frontier*. The Children's Partnership.

Liu, Z. & Meng, Q. (2007). *Usability and Digital Divide: A Study on Usability Issues for Off-Farm Workers in China*. Available from: http://mikeb.inta.gatech.edu/UCDandIDWorkshop/papers/liu.pdf. Last access date: 04 May 2010

Martindale, L. (2002). *Bridging the Digital Divide in South Africa*. Available from: http://www.linuxjournal.com/article/5966. Last access date: 03 May 2010

Mitra, S. & Rana, V. (2001). Children and the Internet: Experiments with Minimally Invasive Education in India. *British Journal of Educational Technology,* 32(2): 221-232.

Nielsen, I., Smyth, R. & Zhai, Q. (2010). Subjective Well-Being of China's Off-Farm Migrants. *Journal of Happiness Studies,* 11(3): 315-333.

Nielsen, J. (2003). *Usability 101: Introduction to Usability*. Available from: http://www.useit.com/alertbox/20030825.html. Last access date: 20 July 2009

Nielsen, J. (2006). *Digital Divide: The Three Stages*. Available from: http://www.useit.com/alertbox/digital-divide.html. Last access date: 30 April 2010

Norman, D. A. (2001). *The Design of Everyday Things*. London: MIT Press.

Preece, J., Rogers, Y. & Sharp, H. (2007). *Interaction Design: Beyond Human-Computer Interaction*. Chichester: John Wiley & Sons Ltd.

Shneiderman, B. (2000). Universal Usability: Pushing human-computer research to empower every citizen. *Communications of the ACM,* 43(5): 84-91.

Shneiderman, B. (2001). Bridging the Digital Divide with Universal Usability. *Interactions,* 8(2): 11-15.

Statistics South Africa (2001). *Census 2001: Key Results*. Available from: http://www.statssa.gov.za/census01/HTML/Key%20results_files/Key%20results.pdf. Last access date: 03 May 2010

Statistics South Africa (2007). *Community Survey, 2007*. Available from: http://www.statssa.gov.za/publications/P0301/P0301.pdf. Last access date: 11 August 2010

Wilson, E. J. (2006). *The Information Revolution and Developing Countries*. Cambridge: MIT Press.