

The Design of Application-Specific Heuristics for the Usability Evaluation of the Digital Doorway

Funmi Adebesein¹, Paula Kotzé²

^{1,2}CSIR Meraka Institute Pretoria and ^{1,2}School of ICT, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa

ABSTRACT

The Digital Doorway (DD) is a joint initiative between the South African Department of Science and Technology (DST) and the Meraka Institute of the Council for Scientific and Industrial Research (CSIR). The DD is a non-standard computer system deployed amongst underprivileged communities in South Africa with the objective to promote computer literacy. Since its inception, there has been no usability or accessibility evaluation of the software installed on the DD, mainly due to lack of usability engineering or interaction design expertise within the development team. The goal of the research presented in this paper was to design a solution to this problem by developing a suitable instrument that could guide DD application developers in the design and development of more usable DD software and interfaces. Design research was used as a research methodology. We first investigated the applicability of the standard usability and accessibility evaluation methods for evaluating the software installed on the DD. During the first cycle of design research, we established that a heuristic-like evaluation method would be an appropriate method for evaluating the usability and direct accessibility support provided by the DD. During a second cycle of design research, embedded in the first, we also developed a set of multi-category heuristics as the ‘instrument’ that could guide the developers during design of applications as well as in the first-level (formative) evaluation thereof. To verify the heuristics, we conducted a usability evaluation of the DD and triangulated the results with a direct field observation at a natural environment of DD use, together with user-administered questionnaires.

CATEGORIES AND SUBJECT DESCRIPTORS

H.5.2 [Information Interfaces and Presentation]: User Interfaces – *evaluation/methodology, User-centered design.*

KEYWORDS

Design research, usability, accessibility, heuristic evaluation, Digital Doorway, digital divide.

1. INTRODUCTION

Despite the huge impact on people throughout the world, many people remain excluded from the potential economic and social benefits of information and communication technologies (ICTs). This phenomenon is typically referred to as the digital divide, which denotes the gap between the information technology ‘haves’ and the ‘have-nots’ [7, 9]. With the ever increasing role of ICTs in the lives of the ‘haves’, the divide is widening. The digital divide is a multi-dimensional phenomenon that extends beyond physical access, or the inability to gain access to ICT devices. Indeed, other factors may more potently contribute to the divide. These include [58]:

- Peoples’ inability to determine information needs, find the information, process and evaluate the information for its appropriateness, and utilize it in a meaningful way (cognitive access).
- The absence of relevant content (content access).
- The problem of complex interfaces (design access). Design access includes both usability and accessibility. Lack of design access is termed the ‘usability divide’ by Nielsen [38].

Various authors have shown the importance of improving usability and accessibility to narrowing the digital divide [10, 38, 58]. In this paper we focus on the design access aspect of

the digital divide, with special attention on usability and, to a lesser degree, on accessibility. Usability is commonly 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” [30]. Without usability, potential users cannot utilize content that might be beneficial and may abandon the system altogether [5, 37]. Accessibility is the design of applications that are “perceivable, operable and understandable for people with a wide range of abilities” [26]. There is a global trend towards enabling access to electronic information by users with a wide range of abilities, to ensure compliance with regulatory/legal requirements and to increase the market share for products and services [25].

The Digital Doorway (DD) project is one of many initiatives currently underway around the world to address the digital divide. It was established by the Department of Science and Technology (DST) of South Africa and the Meraka Institute of the Council for Science and Industrial Research (CSIR), and first deployed in the rural community of Cwili in South Africa’s Eastern Cape Province in 2002. In the decade since, more than 200 DDs have been deployed around South Africa. The DD is a non-standard computer system installed in publicly accessible locations to promote computer literacy and narrow the digital divide [24]. The system offers “a walk-up and use”, custom-designed computer kiosk comprising multiple terminals that enable several users access simultaneously (see Figure 1). Each terminal has a robust housing and metal keyboard with

¹ Email: fadebesin@csir.co.za

² Email: paula.kotze@meraka.org.za



Figure 1. Users at Digital Doorway terminals [34]

reinforced touchpad for input, to minimize vandalism. The system runs pre-loaded software applications and content on the Ubuntu Linux operating system [24].

The DD project has focused, mainly, on providing physical computer access by developing and deploying hardware in locations, such as schools, police stations and community centres, serving otherwise digitally-disenfranchised people. The systems were deployed without conducting any usability evaluation or applicability tests on the software, largely because the team's expertise lie in hardware and software engineering, rather than interaction design or usability engineering. This has meant that widely accepted principles of usability and accessibility were not drawn upon in designing and evaluating applications intended for use on the DD.

Indeed, the hardware in its current form does not conform to many accessibility principles; for instance it cannot support the use of assistive devices, such as screen readers for the visually impaired. On the other hand, within the context of the software, we can evaluate the level of direct accessibility support offered by the DD. Direct accessibility refers to redundancies, built-in to applications, to enable as many people as possible to use the system without modifications [55]; and thus is beneficial to users with or without disabilities. For example, providing text-based feedback, in addition to audio feedback, will benefit a user working in a noisy environment (situational limitation) and enable access to such information by a user who is hard of hearing.

Our research had two primary objectives. Firstly, we wanted to develop an instrument that could guide DD team members and contract developers in the future, in designing and evaluating applications. Within this objective we aimed to develop a set of multi-category usability heuristics that could be used in early or formative evaluation during design and prior to evaluation with end-users. Secondly, we sought to validate the effectiveness of the instrument. We also decided to include heuristics for assessing the direct accessibility support provided by the DD because of the crucial role of accessibility to bridging the digital divide. Thus, in this paper we report on the processes we used to determine appropriate usability and direct accessibility evaluation methods for the DD, and develop and validate the multi-category heuristics instrument.

A subsidiary objective emerged as we started to develop the evaluation instrument. This was the need to demonstrate the value of usability and accessibility evaluation to the developers and DST. Hence, we decided to conduct a usability evaluation on a selection of current interfaces and applications installed on the DD. While this may motivate future research on developing heuristics that could be used in the design of the hardware, at

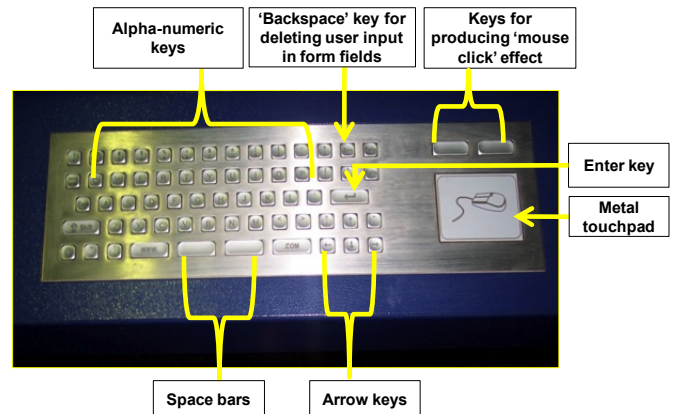


Figure 2. Digital Doorway metal keyboard with touchpad

this stage we focus on software within the context of the current hardware configurations only.

We begin by briefly overviewing the DD and the applications to which we later refer. Next, in section 3, we summarize our research design; then, in section 4, we discuss how we decided on a suitable instrument for evaluating the DD. Section 4 is closely linked to section 5, the design of application-specific heuristics for evaluating the DD. Then, we discuss how we validated the set of heuristics in section 6 and, in section 7, we reflect on applying the instrument we developed.

2. THE DIGITAL DOORWAY

The DD was developed for users with little or no computer literacy from impoverished communities around South Africa. Although the computers are targeted at users of all ages, the majority of DD users are children and young adults [21, 24]. A 2010 study by Gush and De Villiers [23] on overall usage patterns of the DD revealed that the majority of logins were by unregistered users (76%), while registered users accounted for only 24% of the total logins. Only demographic information of registered users can be collected. The ratio between the male/female registered users was 77:23. DDs are used by both single and groups of users, and sometimes as many as twelve users congregate around the terminals. This can allow for individual and peer-assisted learning [49].

The primary input device of the DD is the non-standard keyboard (Figure 2). The keyboard provides alpha-numeric keys but no special function keys. Also, there is a reinforced metal touchpad, which acts as a pointing device.

The main output device of the DD is a robust liquid crystal display (LCD) screen (see Figure 1), covered with plexi-glass to prevent vandalism [23]. DDs are not attached to printers; hence users must stand in front of the computer and read content from the screen. Users cannot save information 'downloaded' from the DD to portable storage devices because there are no 'ports' to which such devices can be attached. Hence, users have to write content onto paper if they wish to take it with them. In addition, users cannot upload content onto the DD.

The DD provides extensive content most of which is open source or third-party applications [22], with additional content developed in-house by contract and visiting developers. This includes: educational games; reference materials, in Wikipedia-like documents; OpenOffice suites; The "Mindset applications", a South African curriculum-based educational program; interactive science simulations; and audio books. Applications are preloaded and their content updated on a regular basis [24]. Given the scale of this content and that it includes third-party

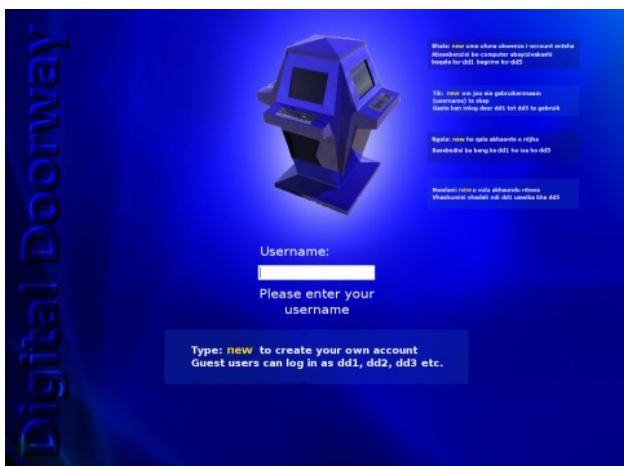


Figure 3. Digital Doorway login screen

materials, which the DD team has no control in developing, we selected content to assess that might best influence positive future development efforts by the DD project team. That is, we chose content developed in-house for the DD during the validation phase of the multi-category heuristics. This encompasses: the login screen, the registration form for creating a new user account, the main desktop, and three educational games.

2.1 Digital Doorway Login Screen

The login screen (Figure 3) is the first interface between users and the DD. The primary language of instruction on the login screen is English, with equivalent instruction in four other South African languages (IsiXhosa, Afrikaans, Sotho, and Venda). Both registered and guest users can access the system. Registered users log in by providing their username and password in succession. When the user types 'new' in the username textbox the registration form, for creating a user account, is activated. Guest users can access content through a one-step process by typing 'dd1' in the username textbox.

2.2 Digital Doorway Registration Form

To create a new account, users fill a simple electronic form (Figure 4). Components of the form are organized into two main groups: 'Personal Details' and 'User Details'. The Personal Details groups user information such as name, age, gender, home language, and preferred language while the User

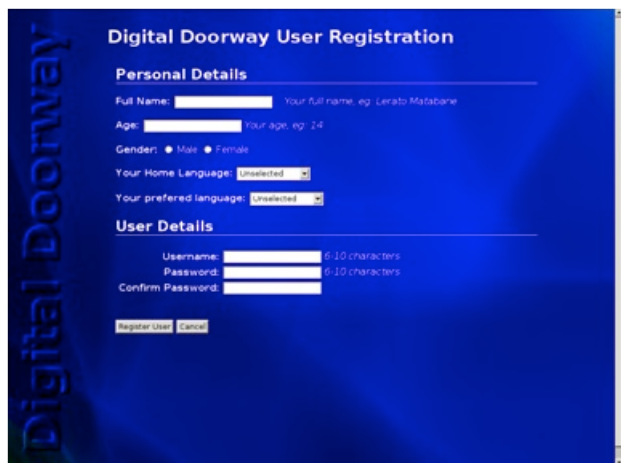


Figure 4. Digital Doorway registration form

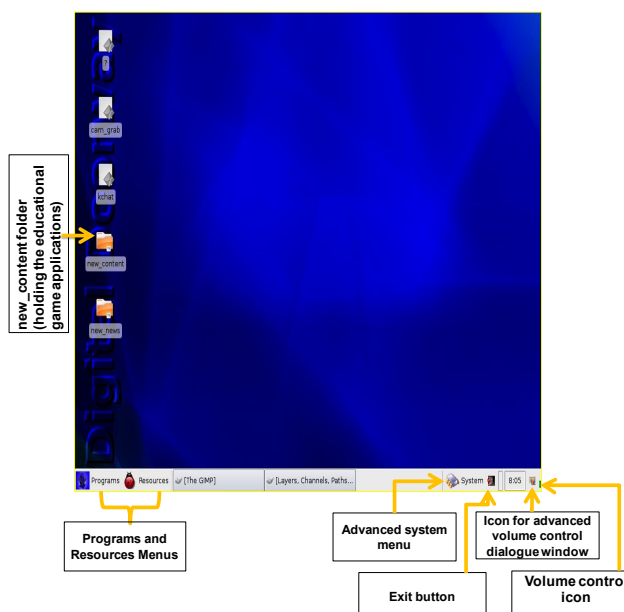


Figure 5. Digital Doorway Desktop

Details groups the username and password. The form hints the type of data expected in certain fields, for example the password field. All the data fields in the form must be filled, although this is not explicitly stated on the form. The end of data entry is signalled by clicking on the <Register User> button. A click on the <Cancel> button closes the form, without a request for a confirmation from the user.

Information provided by users is automatically transferred to a central server on a daily basis. Currently this information is only used by the DD team to gather demographic data and application usage patterns [23]. The information provided is not used to customize the applications. For example, the language fields do not affect the default language setting in *OpenSpell* (see section 2.5).

2.3 Digital Doorway Desktop

Users who successfully log in to the DD system can access content 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 through a click on a volume control 'icon', represented by a green slider bar at the extreme right of the taskbar, illustrated in Figure 5, or through an advanced volume control dialogue window. Users can log out of the system by clicking on an exit 'button' (designated by a red right pointing arrow ⇒) or from the advanced system menu (located on the taskbar).

2.4 What-What Mzansi

What-What Mzansi (Figure 6) is an educational quiz game, in the form of yes/no questions, developed to provide content relevant to the South African environment. The game is available only in the English language. When a user clicks on the <?> icon at the top right corner the application provides context-specific instructions, while clicking on the <X> icon next to it closes the application. The game presents questions which the user can answer when he/she selects the <play> menu. There are two levels of difficulty from which a player can choose: <Easy> and <Advanced>. When the user selects a difficulty level, a voice welcomes him/her and ask a set of questions, which can be answered by clicking on <Yes> or <No> on the game interface, or by pressing the arrow keys on



Figure 6. Interface of What-What Mzansi

the metal keyboard. Each session, comprising a number of questions, lasts 60 seconds. Scoring is based on the player's speed in answering the question and ranges from 2, for the slowest response, to 10 for the fastest, provided the answer is correct. The number of questions asked in one session also depends on the speed of the player. The <high-scores> menu lists the scores of the top-ten registered players. When the user clicks the <about> menu, the application presents information on the DD project and its achievements, together with information about the game developers.

2.5 OpenSpell

OpenSpell is an educational spelling game. It is available in all of the eleven South African official languages and provides three levels of difficulty designated with *, **, ***. When the user launches the application, the default language is English and the difficulty level is the simplest form. The user can select another language which then changes the interface and content to the language of choice; however, it reverts back to English when the application is restarted.

The interface includes an onscreen keyboard for providing input in spelling exercises. There are three game mechanics. When the user clicks the <say> option, (Figure 7) the application provides a series of pictures of words which the user learns their spelling. For each word, a voice in the chosen language speaks out each letter, together with its pronunciation. When the user selects the <guess> option a hangman word guessing game is presented; but when the user selects the <spell> option the game mechanic is spelling exercises. The user spells the words by clicking letters of the onscreen keyboard and has two attempts, after which the correct answer is presented to him/her.

2.6 Themba's Journey

Themba's Journey, shown in Figure 8, is a life skills program in the form of a narrative story of Themba, a young man who travels from his village to the city of Johannesburg in search of a job opportunity. The main story is narrated within the <PLAY> environment. At strategic places, Themba reaches crossroads where the user has to make decisions on his behalf

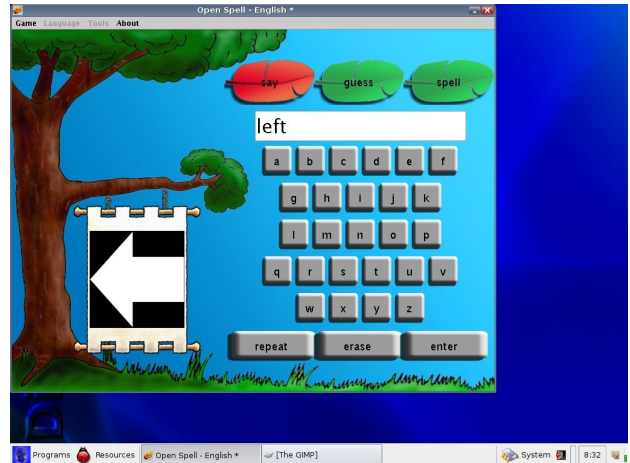


Figure 7. OpenSpell Interface

on a course of action, e.g. to take drugs or not. Each action has a positive or negative consequence. The game is available in IsiXhosa (one of eleven South African official languages) and English. The default language is IsiXhosa, available in text and spoken words. A text equivalent of the narration is provided in English, but it can only be accessed when the user hovers the pointer on speech bubbles. The <HELP> menu contains the navigation and game instructions. The <EXIT> menu closes the application.

3. RESEARCH DESIGN AND METHODOLOGY

In this section we summarize the processes we used to determine appropriate usability and direct accessibility evaluation methods for the DD and design the heuristic-like guidelines. As we discussed in the introduction we needed the heuristics to assist DD application developers who are not necessarily usability experts, to design, and formatively evaluate, more usable applications.

Our study is based on the design research paradigm as defined by Vaishnavi and Kuechler [55]. Design research is concerned with the expression of man-made things, how these things perform their jobs, and how they work [6]. It involves analyzing the performance of designed artefacts, such as a computer system interface, with the aim of better understanding and improving the artefact. The design research paradigm provided us with a framework to undertake the iterative

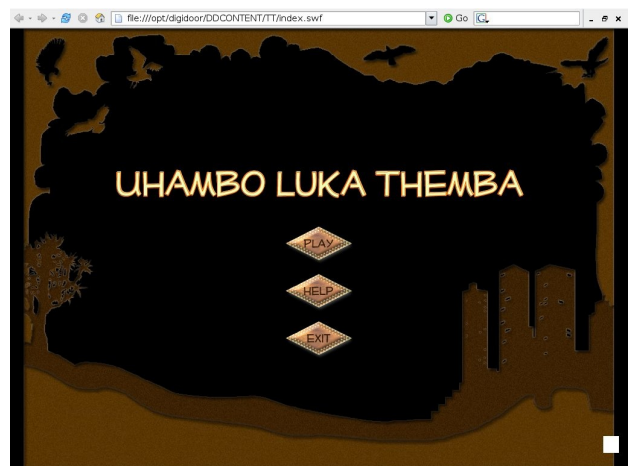


Figure 8. Interface of Themba's Journey

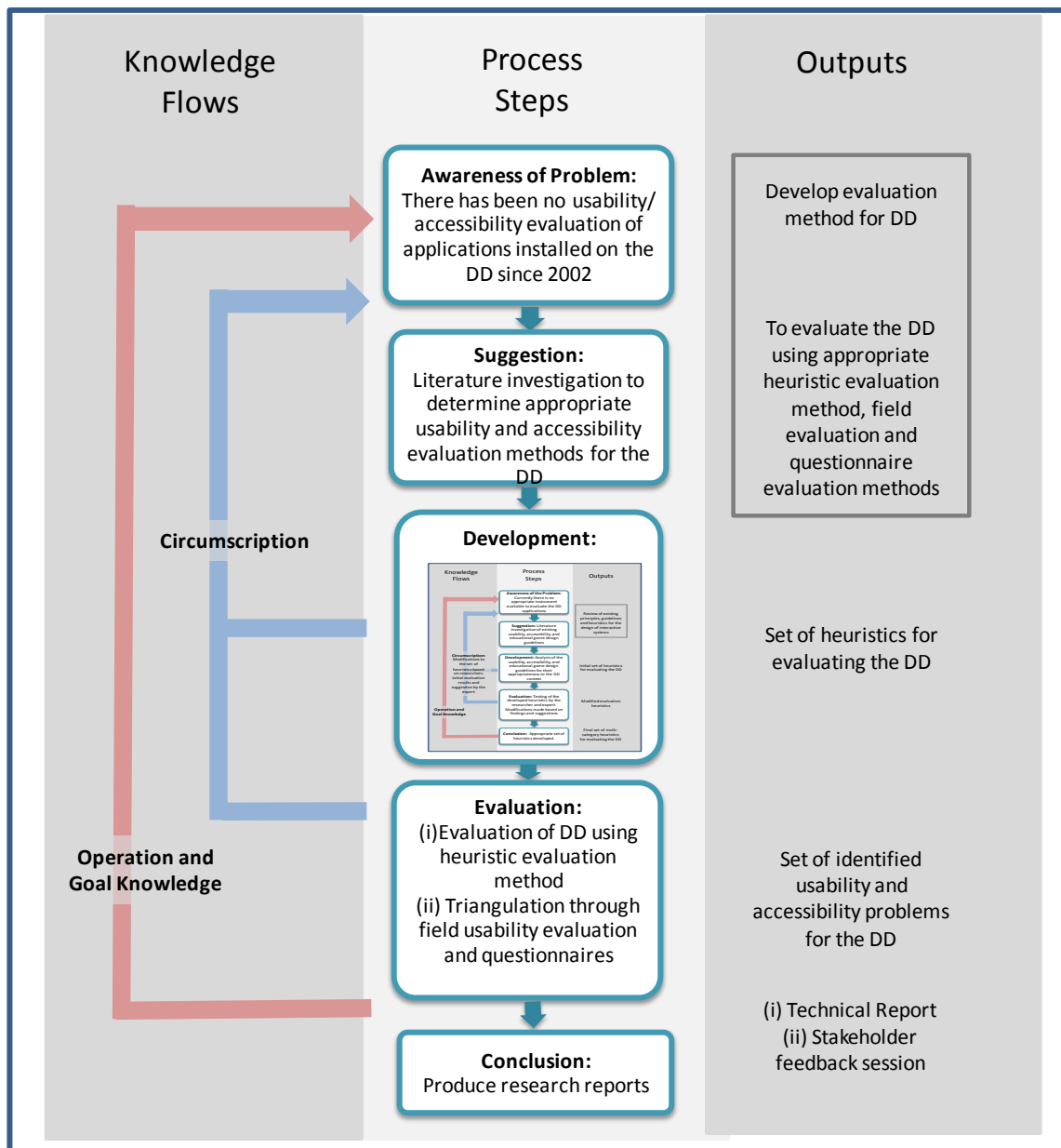


Figure 9. Design research process for designing an appropriate evaluation instrument for the Digital Doorway

activities necessary to develop an instrument that could guide the design of more usable applications for the DD

Our choice of the design research paradigm, as opposed to the positivist paradigm, was motivated by the goal-directed nature of the study, a fundamental requirement of design theory [56]. Goals involving the production of an artefact has no meaning in the realm of natural science (positivistic) theory save for those by the theorists. Likewise, in the social science (interpretivistic) theories, the purpose is not the achievement of such goals, but to explain their existence or predict the outcomes of specific goals [56].

In our design research, our goals were:

1. To design an instrument that could guide the DD team in the development of more usable applications for the DD.
2. To validate the effectiveness of the instrument by conducting an evaluation of a selection of interfaces and applications on the DD.

Design research generally consists of five iterative phases starting with an awareness phase and ending with the conclusion phase [55]:

1. *Awareness of problem*, where the need for the new research is established.
2. *Suggestion*, where new functionalities are envisioned.
3. *Development*, where the artefact is created based on a tentative design.
4. *Evaluation*, where the designed artefact is evaluated for any deviation from what is envisaged.
5. *Conclusion*, which represents the final phase of the current research effort and culminates in the development of an artefact that is judged as ‘good enough’.

We undertook two cycles of design research phases, one (Figure 10) nested inside the other (Figure 9). In the first outer cycle, we needed to determine appropriate method(s) for evaluating the DD (Figure 9.) This process identified a

heuristic-like evaluation method as the most appropriate method to evaluate the usability and direct accessibility support provided in the DD. In the second and inner cycle we developed specific heuristics for evaluating the DD (Figure 10). Next we discuss the awareness and suggestion phases for the design of the suitable instrument for evaluating the DD.

4. DESIGNING AN EVALUATION INSTRUMENT

4.1 Awareness of the Problem Phase

Becoming *aware* of a problem constitutes the first phase of the design research paradigm (see Figure 9). Thus, we marked the start of the entire process by realizing that although the DD has been operational since 2002 there has been no evaluation of the installed software. Prior to our study, the DD project concentrated efforts on providing physical access to computer systems rather than on the usability or accessibility of the applications installed on the system. However, without usability, there is a risk that user communities might not benefit from content as they can't get to it. Further, a complex interface can lead to users perceiving themselves as unable to use a computer, or low computer self-efficacy [8], which can in turn result in people abandoning the DD altogether. We sought to produce an instrument for use by people who do not have expertise in usability and accessibility, especially in the context of the digital divide [10, 38, 58], during development and evaluation. At the same time, we recognize that while the DD development team are not usability experts, or experienced in interaction design and usability engineering, they are highly experienced software developers.

4.2 Suggestion Phase

In the *suggestion* phase we investigated literature on standard usability and accessibility evaluation methods with a view to determining which of these would suit evaluating the software developed or installed on the DD. Our selection was influenced by several initial criteria. Firstly, the method(s) should be reasonably easy to use by software developers who are non-usability and non-accessibility experts and fit the context of use of the DD. Secondly, method(s) needed to suit physical and logistical constraints imposed by the DD. Thirdly, since the DD aims to address the digital divide we needed to involve end-users in real context of use in the evaluation. Finally, methods needed to be practical within available resources. In the remainder of this section we summarize the primary usability and accessibility evaluation methods we considered and conclude by recommending the most suitable.

4.2.1 Methods for Evaluating Usability

Usability evaluation methods (UEMs) can be classified into two groups. The first group require expert evaluators:

- *Heuristic evaluation* involves multiple evaluators independently assessing an interface using a set of heuristics to identify potential usability problems [36]. The evaluation can be used for both formative and summative purposes but may require developing application-specific heuristics.
- *Cognitive walkthrough* is an inspection method derived from cognitive science, where expert evaluators with cognitive theoretic skills evaluate a given interface by stepping-through a set of tasks [13, 16, 43, 57]. The focus of cognitive walkthrough is on the learnability quality of the application without consideration for other measures of usability, for example, efficiency [57].

The second set of evaluation methods requires user involvement:

- *User observations* involve observing end-users while they interact with the target application in the environment of application usage, within a controlled setting or, by reviewing recordings of user activities. Observation in the context of usage can provide insight into usability problems that might be difficult to obtain through other evaluation methods. User observations are useful early during the design stage as it enables an understanding of the context of users' tasks, as well as during summative evaluation of a fully functional application [16, 17].
- *Usability testing* involves measuring the performance of end-users while they execute predefined tasks [5, 43, 45]. It uses quantitative usability metrics, such as time spent on tasks, the number of errors, time taken to recover from errors, as well as qualitative data, for example, verbal expression of frustration or satisfaction are recorded for later review [5, 35]. Usability testing can be complimented with the think-aloud method [5, 16, 43] for better insight into the rationale for user actions, and eye tracking [18, 42] to pinpoint specific components with usability problems.
- *Questionnaires* are well-established in human-computer interaction (HCI) as ways to gather information on users' attitudes and preferences and evaluate products [16, 40, 43]. They can be used as an evaluation method alone or combined with other evaluation methods. For example, questionnaires can be used after usability testing to assess participants' opinions on the application just tested [17, 43].
- *Interviews* are a data collection method that can be employed at early design stages to establish users' requirements for an envisioned application, or as a summative evaluation of a system that is already in use. Interviews can be structured or unstructured, depending on the goal of the evaluation [43].

4.2.2 Methods for Evaluating Direct Accessibility

Direct accessibility evaluation can be combined with some of the usability evaluation methods we have just discussed (section 4.2.1). Now we turn to the accessibility evaluation methods that we considered for our study.

Greeff and Kotzé [20] and Henry [26] recommend the following accessibility evaluation methods:

- *Standards review* involves assessing an interface for compliance against predefined standards and is typically employed prior to the involvement of users with disabilities in the evaluation process. These standards might be internally specified within an organization or externally stipulated by a national or international body, for example, the World Wide Web consortium's (W3C) web content accessibility guidelines (WCAG 1.0) [59]. Standards review requires the evaluator to be competent in the specific standards against which the interface is being evaluated.
- *Heuristic evaluation*, in the context of accessibility, requires using heuristics aimed to address specific accessibility requirements. Usability and accessibility evaluations can be combined in a heuristic evaluation.
- *Design walkthrough* involves an evaluator imitating the behaviour of a typical user in using the system guided by personas and scenarios. Specifically, such personas will include characteristics of potential users, such as colour blindness. The scenario describes how the persona completes a task using the target application with some

form of adaptive strategy or an assistive technology, for example, a screen reader. Design walkthroughs are useful at early stages of development to uncover potential accessibility problems.

- *Screening techniques* aim to simulate, during an evaluation, aspects of the constraints users with specific disabilities contend with while using interactive systems. Evaluators aim to experience some of the constraints by modifying one or more of their sensory and/or motor capabilities before interacting with an interface; for example, by wearing low vision glasses to limit vision or a thick glove to reduce hand dexterity.
- Standard *usability testing* can be modified to incorporate accessibility evaluation by involving users with disabilities as test participants. The number and characteristics of participants that are included in the evaluation will depend on the disability category to which the application is targeted.
- *Accessibility evaluation tools* include software that can automatically detect defined problems against accessibility standards and guidelines. For instance, tools such as Total Validator [52] can be used to evaluate web pages and elements of software for compliance with standards like WCAG 1.0 [59]. Humans are required to establish the nature of the identified problem, for example, a software tool can indicate a missing *alt text* for an image but cannot determine whether an existing *alt text* conveys the same information as the image it represents.

In deciding on appropriate methods to use we were also guided by the *lightweight accessibility evaluation methodology* by Greeff and Kotzé [20]. The Greeff and Kotzé [20] methodology refers to developing in-house context-specific guidelines, which suit the specific development environment and could assist developers of web pages, who are not accessibility experts, in developing more accessible pages. The method proposed three phases for evaluating web pages. It commences with an initial evaluation with automated tools to check the conformance of a web page with specific accessibility guidelines, the readability of texts, colour contrasts, and its exposure to screen readers. Next, it tests with representative users in different disability categories to ensure that these groups will be able to use the web site. Finally, it involves developing in-house context-specific guidelines, based on findings from the first two phases.

4.2.3 Suggested Usability and Accessibility Evaluation Methods for the Digital Doorway

Our choice of evaluation method(s) needed to suit developers of DD software, who typically have little or no experience in interaction design or usability engineering and, as we explained in Section 4.2, be suited to the DD. Our assessment of the various usability and accessibility evaluation methods revealed that several were unsuitable. For instance, the DD is not suited to controlled usability studies in a state-of-the-art usability testing laboratory, even though it is on hand and in-house expertise for running such studies is available. Firstly, we were unable to physically move the DD to the usability laboratory. Secondly, observation and logging software in the usability laboratory are only compatible with the Windows operating system, while the applications on the DD run on the Ubuntu Linux operating system. Likewise, the use of automatic

software tools as recommended by Greeff and Kotzé [20] and Henry [26] is not adequate because most of the available tools are targeted at evaluating web pages and applications running on the Windows operating system. Cognitive walkthrough methods are also impractical and inadequate for evaluating the DD. While such walkthroughs suit summative evaluations, they tend to focus only on the learnability aspects of an application and also require evaluators to possess skills in cognitive theory. The design walkthrough and screening technique for accessibility evaluation proposed by Henry [26] are not applicable to our study, since the DD does not allow for the use of assistive devices in its current configuration.

We found that the heuristic evaluation method, direct user observations, questionnaires and interviews were most suited to evaluating the DD. A heuristic-based type of evaluation can be used by developers to assist them in the design and first level evaluation of applications and software. Further, it is suited to the constraints of the DD for evaluating both the usability and direct accessibility support. Questionnaires and interviews are targeted at end-users, and observing users in the context of use provides valuable insight into use in the under-resourced communities into which DD is deployed. Thus, we combined these methods in the design and evaluation of the multi-category heuristics instrument. As we describe in Section 6 this combination required different validation endeavours. Expert evaluators conducted a usability evaluation of the DD to validate the effectiveness of the set of heuristics we developed. Then, we triangulated direct observations and questionnaires in a field study at a school where the context of DD usage was retained. We also adopted Greeff and Kotzé's [20] recommendation such that the set of application-specific heuristics would provide guidance for DD applications developers. This necessitated a higher granularity, requiring a more detailed set of heuristics.

The heuristic evaluation method can be used by experts and non-experts [12], although expert evaluators typically produce better results [15, 27]. Further, the method is widely accepted as relatively easy and cost-effective in evaluating the usability and accessibility of interactive systems' interfaces [12, 26]. It can uncover a large number of potential usability and accessibility problems without the involvement of end-users.

Heuristic evaluation is not without drawbacks. It requires multiple evaluators to detect substantial numbers of potential problems, especially when the heuristics are highly generic. We found the set of heuristics developed by Nielsen [34] to be inadequate for our study. These heuristics require usability expert evaluators and are too general to serve as guidelines in developing software for a specific target use group. Indeed, Nielsen (as cited in Rogers et al. [44]), acknowledges the inapplicability of this set to contemporary applications, and suggests developing application-specific heuristics when necessary. A wide number of application-specific or domain heuristics have emerged, including the Pierotti's Xerox Heuristic Evaluation Checklist [41]; Travis's web usability guidelines [53]; Desurvire et al.'s heuristics for evaluating the playability of games [14]; and, Sim et al.'s heuristics for evaluating the usability of computer-assisted assessment (CAA) applications [48]. One of the most prominent sets of heuristics, prior to Nielsen coining the term 'heuristic evaluation', is most probably the 1986 set of guidelines for designing interface software by Smith and Mosier developed for the United States Air Force Systems Command [50].

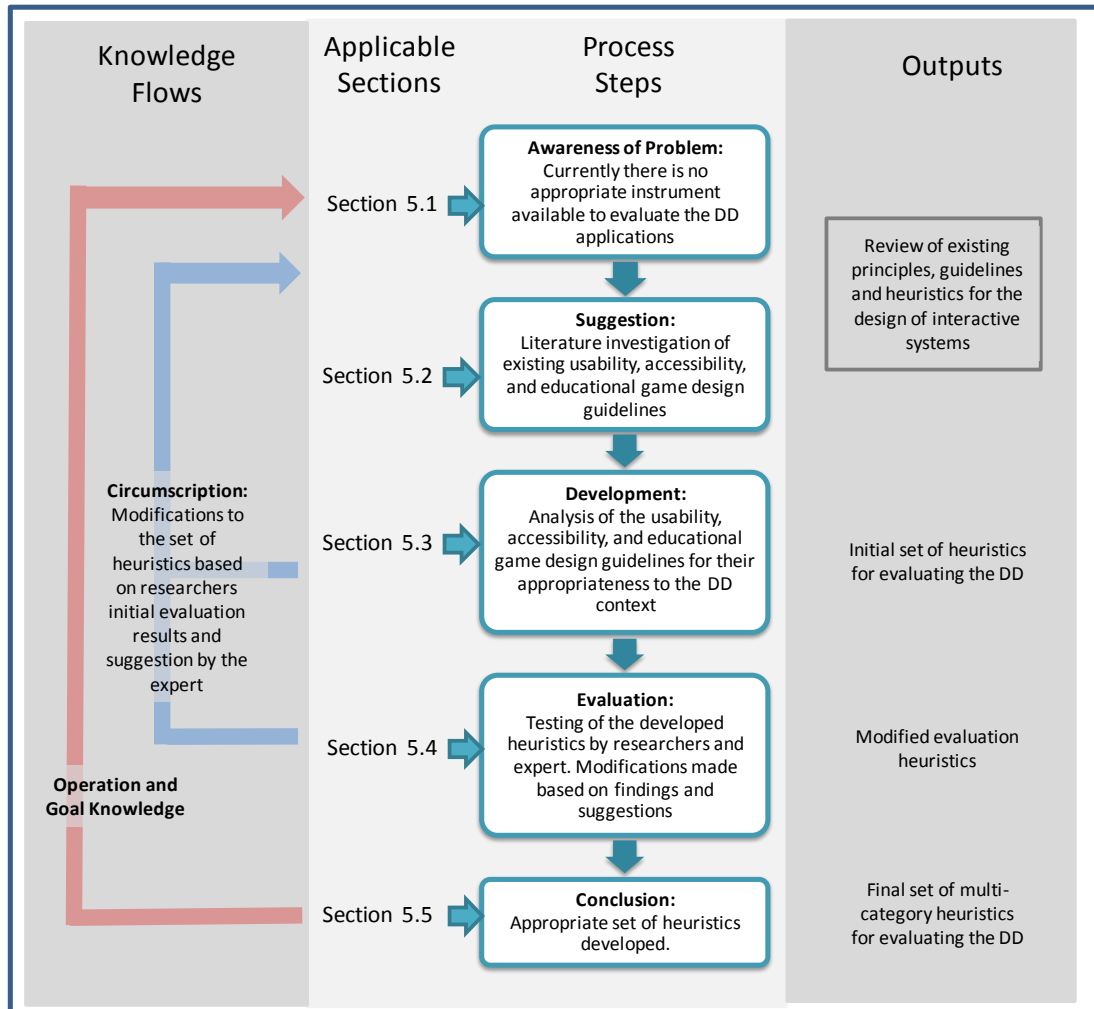


Figure 10. Design research process to develop application-specific heuristics for the Digital Doorway

The specific requirements of each context influence the number of heuristics required in the set and the heuristics are typically grouped into several categories to cover the various evaluation objectives. Thus, application-specific heuristics are generally much larger in number than the original set of ten heuristics proposed by Nielsen [36].

Once we identified heuristic evaluation as the primary evaluation method and the need for multi-category application-specific heuristics, we entered the development phase of the design research process. The development phase addresses the second cycle of design research which, for our study, develops the heuristics for use in the suggested evaluation methods as we discuss next.

5. DESIGNING APPLICATION-SPECIFIC HEURISTICS

The development phase of our design research involved a second, and inner, cycle to develop the set of multi-category heuristics for guiding the design and evaluation of the DD (Figure 10).

5.1 Awareness of problem Phase

We identified a heuristic-based type of evaluation as an appropriate method to evaluate the usability and direct accessibility support, in the suggestion phase of the outer design research cycle (see section 4.2.3). We also identified that we

needed a set of application-specific heuristics for the DD, focusing on usability and direct accessibility. This catalyzed the awareness phase of the inner design research cycle.

5.2 Suggestion Phase

The real value of the heuristic evaluation method lies in the use of appropriate heuristics. As we explained in section 2, the DD is aimed at users with little or no computer literacy. Hence, it was essential to consider general usability principles, such as those relating to intuitive interfaces and the provision of adequate feedback. It was also important to consider guidelines with a specific focus on direct accessibility, which is also required to bridge the digital divide. Further, general usability guidelines do not cover the specific usability of educational applications, which constitute both the majority of applications developed in-house for the DD and the three we selected for use in the validation exercise. We accounted for all of these issues in the suggestion phase as we reviewed, in-depth, various usability, accessibility and educational game usability design guidelines.

5.3 Development Phase

In the development phase we actually developed the multi-category heuristics. We started by reviewing a large number of relevant guidelines and analyzed them for their applicability to our context.

We reviewed Dix et al.'s principles of usable interface design [16]; Gelderblom's guidelines for the design of children's technology [19]; Mayhew's guidelines for the design of form-fill interfaces [33]; Nielsen's heuristics [36]; Rogers et al.'s usability principles [43]; Norman's design principles [39]; and Shneiderman's golden rules for interface design [47].

To address direct accessibility requirements, we examined the seven principles of universal design [51]; the United States' standards for electronic information accessibility (Section 508) [54]; IBM software accessibility checklists [29]; and, WCAG 1.0. [59] (WCAG 2.0 was still under development when the heuristics were designed). We also reviewed guidelines that relate specifically to educational applications since these comprise the majority of the DD applications developed. This review included the game-specific guidelines proposed by Alessi and Trollip [3], Desurvire et al. [14], Malone [31, 32] and Shelley [46].

Our decision to include or exclude specific design principles and guidelines was guided by the nature of the interfaces and applications to be evaluated, the types of users the DD targets, and the typical environment of DD usage. For example, we found that the principle of multithreading, as proposed by Dix et al. [16], was not pertinent to the usability of the DD since the system only supports the execution of one task at a time. We grouped relevant heuristics into four categories: general usability; form usability; direct accessibility; and, educational game usability heuristics.

5.4 Evaluation Phase

To validate the effectiveness of our heuristics, we used an iterative process of design-evaluate-redesign that commenced with an initial heuristic evaluation on the selected interfaces and applications. We then matched the problems that we identified to the heuristics and adapted the heuristics to cover the type of problem identified when we were unable to match the problem to a heuristic.

After the initial evaluation, an expert with usability and accessibility experience undertook another round of evaluation. This exercise aimed at assessing the completeness of and the terminology we used in the heuristic set. We modified the heuristics based on the expert's recommendations to produce the set of multi-category heuristics that was used in a formal evaluation.

To ensure that the heuristics provide adequate coverage for the DD context, and meet the dual objectives of evaluating the usability and the direct accessibility support provided in the DD, a team of five experts used the heuristics to evaluate the selected interfaces and applications on the DD. This was followed by a field evaluation with users at a natural usage environment (see section 6). The evaluation processes did not uncover inadequacies in the terms and language used to describe the heuristics.

Furthermore, Sim et al. [48] defined criteria that can be used to assess the effectiveness of application-specific heuristics, in addition to the number of usability problems identified through the use of such heuristics. These criteria are *correctness of terminology*, and *coverage and thoroughness*. Correctness relates to the use of appropriate terminology to describe the heuristics, while coverage and thoroughness refer to the completeness of the heuristics set. We also used the criteria in a final assessment of the heuristics.

5.5 Conclusion Phase

In the conclusion phase, we decided the heuristic set was appropriate for evaluating the DD following the iterative process of development, evaluation and modifications. This

Table 1. Profiles of expert evaluators

	Qualification	Rank	Roles at Institution
1	MSc (Information Systems)	Senior Lecturer	Tuition; Postgraduate supervision; Usability evaluations.
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; Organising and conducting usability and accessibility evaluations.
4	PhD, MSc, MEd	Professor	Tuition; Postgraduate supervision; Research; Management and leadership; Usability evaluations.
5	M(Eng): Technology Management	Researcher	Research in the field of voice user interfaces; Conduct of usability and accessibility evaluations.

concluded the inner cycle of design research and resumed the evaluation phase of the outer cycle. This phase, as we explain next, validates the multi-category, application-specific heuristics by evaluating a selection of interfaces and applications on the DD.

6. VALIDATING THE MULTI-CATEGORY HEURISTICS

The evaluation phase (of the outer cycle of design research) involved assessing the efficacy of the developed multi-category heuristics through formal evaluation of the DD. This assessment involved expert evaluators, using the set of application-specific heuristics developed, and end-users, using direct observations and user-administered questionnaires.

6.1 Heuristic Evaluation by Experts

In this section we discuss the formal evaluation of the DD through the heuristic evaluation method using the set of multi-category heuristics developed, as well as the evaluation results of applying this method.

6.1.1 The Heuristic Evaluation Process

During the first phase of the evaluation process, a team of usability/accessibility experts evaluated the DD using the set of multi-category heuristics developed. Due to physical, logistical and work-related constraints, evaluation was conducted at the DD laboratory at the CSIR Meraka Institute over a period of three months [1].

In selecting the evaluators, we accounted for expertise and the optimal number of evaluators [36]. Hence, five evaluators participated in the formal heuristic evaluation, three with usability expertise and the other two had experience in usability and accessibility evaluations (Table 1).

To ensure uniformity in the briefing and allow evaluators sufficient familiarization time, we mailed a comprehensive information document to all the experts in advance. The documentation provided an overview of the interfaces and applications to be evaluated, profile of typical users of the DD, the procedure to be followed during evaluation, as well as the evaluation heuristics to be used (Appendix A). Evaluators

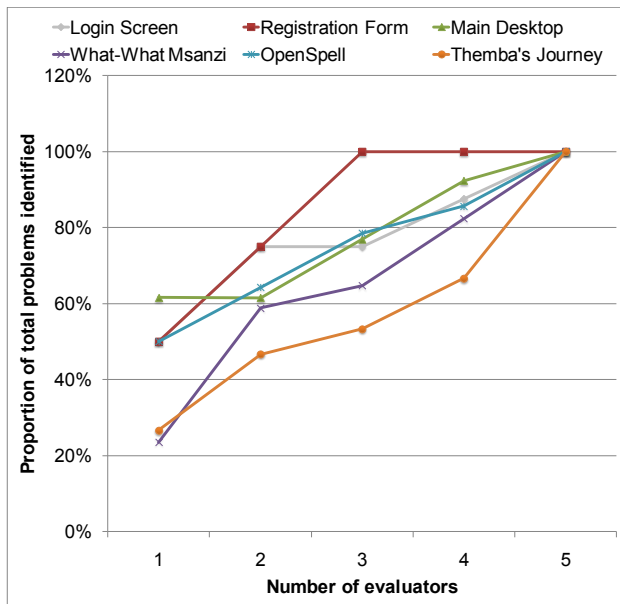


Figure 11. Proportion of problems per interface/application as more evaluators are included

signed an informed consent form, which assured their anonymity and the confidentiality of information provided.

Evaluators performed typical user tasks involving the use of all three educational game applications. Each evaluator went through a two-step process. The first step involved familiarization with the interfaces and applications through a guest login. The second step involved the actual evaluation where evaluators created user accounts before login.

The physical configuration of the DD required evaluators to stand in front of the system while conducting the evaluation. This made it difficult for evaluators to make notes of their findings while conducting the evaluation. To ease their task one of the authors acted as a scribe. This allowed evaluators to focus on the primary task of identifying usability and/or direct accessibility problems based on the heuristics provided. Furthermore, an evaluation report was compiled by the scribe and mailed to evaluators a day after the evaluation. The report was then verified by each evaluator to ascertain that it was a true reflection of the evaluation. In some cases, the verification resulted in modifications to the report. Thus, evaluators did not have to produce evaluation reports from scratch, and we had access to consistently structured reports according to the level of detail required.

6.1.2 The Heuristic Evaluation Results

We aggregated the set of usability and direct accessibility problems identified in the heuristic evaluation by all the experts and consolidated those that were of similar nature. Then, we analyzed the identified problems based on the number of evaluators that detected the specific problems. We also categorized the set of identified problems according to the evaluation heuristics that were violated by any particular problem. Our further analysis was based on the specific interface/application in which problems were located. We utilized descriptive statistics to provide summaries of the complete set of unique usability problems as well as the number of problems identified by each evaluator.

The five expert evaluators between them identified an aggregate of 71 usability and direct accessibility problems. However, only a narrow proportion of the total problems were identified by each evaluator, ranging between 45% and 55%.

Table 2. Number of recorded problems per interface/application

Interface/Application	No of Problems
Login Screen	8
Registration Form	4
Main Desktop	13
What-What Msanzi	17
OpenSpell	14
Themba's Journey	15

This is not unusual, as it is generally difficult for any single evaluator to identify all the potential problems in an interface. For example, one evaluator detected only 24% of the total problems affecting the game application, *What-What Msanzi*. The number of problems identified increased with the number of evaluation sessions conducted and the number of problems detected per interface/application increased with the number of evaluators (Figure 11). This explains why multiple evaluators are required to conduct summative heuristic evaluations.

The least number of problems, that were identified, were located within the registration form, and the highest number occurred in *What-What Msanzi* (see Table 2). However, taking the login screen and the main desktop together we found a total of 21 problems. That is, the combined number of problems within these two interfaces was more than that found in any other interface or application evaluated. This is particularly problematic for access and inclusion. Since these two interfaces are the first contact that all users have with the DD, whether they are registered users or guests, they could affect the majority of users.

Of the 71 aggregated problems identified we found that 61% violated general usability heuristics, 37% were related to game usability heuristics, 21% violated direct accessibility heuristics, and only 4% related to the form usability heuristics. The high number of problems that violated general usability heuristics can impact on achieving one of the goals of the DD project, which is to promote computer literacy through unassisted learning [11]. Applications aimed at promoting unaided-learning should adhere at least to basic usability principles. We expected a relatively high number of problems violating direct accessibility heuristics, since the DD was not designed with accessibility in mind. However, compliance with the direct accessibility heuristics will go a long way towards improving general usability for users with no apparent disability, since they use the system in an environment that imposes some limitations; for example, noisy surroundings and glaring sunshine. Many of the problems violating the game usability heuristics were also general usability problems. Thus, correcting the general usability problems will also address many of the game-related problems.

The background expertise of the evaluators also influenced the number of problems identified per heuristic category. Evaluators identified a similar proportion of problems that violated general usability heuristics; although this was not the same set of problems (see Figure 12). The results illustrate the benefit of including evaluators with accessibility skills. Evaluator 3 and 5 each identified 53% of the total problems relating to the direct accessibility heuristic category, while evaluators 1 and 4, without accessibility evaluation experience, identified the least problems (33%) violating this heuristic category.

We were unable to match two of the problems, detected in the formal evaluation by the team of experts, to any of the heuristics in the set of multi-category heuristic developed. Thus, we generated appropriate additional heuristics to provide for these problems. One of these problems related to the absence of

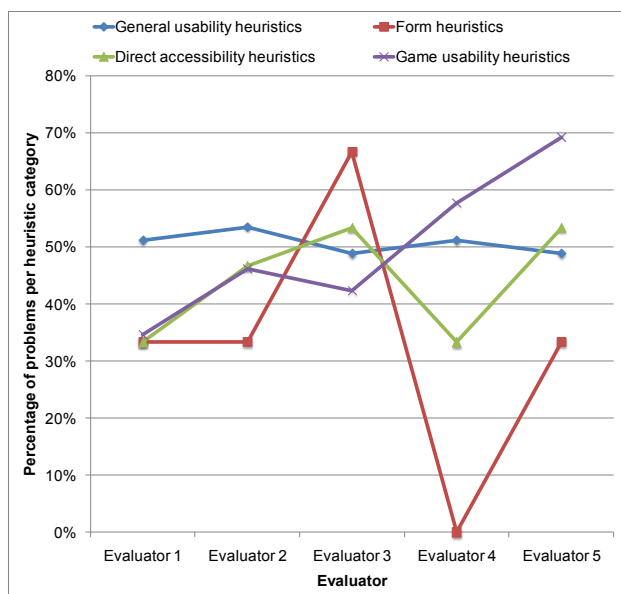


Figure 12. Problems identified by evaluators per heuristic category

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.

Our ability to refine the heuristic sets through the iterative process of design-evaluate-redesign was valuable. Not only did it result in the final set of heuristics (see Appendix A) but it provided insight into correcting the identified usability and direct accessibility problems. The experts’ recommendations were a vital part of correcting the problems. The complete set of usability and direct accessibility problems, together with recommendations is so extensive that we cannot present them here but are documented in Adebisin [2].

6.2 Field Evaluation with End-users

We conducted a field evaluation at a school where the DD is installed to verify the heuristics and their coverage of potential usability and direct accessibility problems. We sought to evaluate the DD from users’ perspectives, as well as triangulate the data obtained from the heuristic evaluation using direct observation and a participant questionnaire. We adopted a deep instead of broad approach by carefully observing a small number of users, rather than compromising the quality of observations by covering a large number of users.

6.2.1 The Field Evaluation Process

We evaluated the DD at a secondary school with a two-pronged primary objective, to determine whether we had omitted any important heuristics and whether the use of the set of multi-

Table 3. Profile of field study participants

Age	Gender	Grade	Application(s) used
17	F	11	OpenSpell
18	F	11	Themba’s Journey
15	M	8	What-What Mzansi
13	M	8	OpenSpell
13	M	8	(Free Exploration) KTuberling and Penguin games
15	F	10	(Free Exploration) Themba’s Journey
16	M	-	What-What Mzansi
14	M	-	Themba’s Journey
15	M	-	(Free Exploration) What-What Mzansi and Four-in-a-row game

category heuristics on its own would be sufficient in evaluating the applications/software installed on the DD.

There were various potential sites for evaluating, such as community halls or police stations, but two factors supported evaluating the DD in a school environment. Firstly, the three applications evaluated were educational games. In addition, secondary schools recorded the most successful usage for DD applications [23].

The DD at the particular school was installed in an open area along one of the school’s corridors. This was to allow easy access to school children, as well as other users living in the area surrounding the school. To address ethical concerns, we obtained formal permission from the school’s principal and written informed consent from the parents/guardians of participants. We conducted evaluation sessions in the afternoons after the official closing hour of the school to reduce disruptions to learning.

Three female and six male learners aged between 13 and 18 years participated in the field study (Table 3). We adapted conventional field observation methods to the context, by asking six participants to undertake pre-defined tasks and letting three participants select any application they wished to interact with. This allowed us to focus on the selected interfaces and applications, while at the same time yielding some insight into the type of applications that are used at the school. We presented pre-defined tasks in text, as illustrated by the sample task list in Figure 13, and also provided verbal explanation of the tasks to participants.

6.2.2 The Field Evaluation Results

We recorded observations on video and in hand-written field notes which we first grouped according to the applications that participants used. We then grouped together similar usability and direct-accessibility problems, using descriptive texts.

Table 4. Number of actual user problems per interface/application

Interface/Application	Number of Problems
Login Screen	6
Registration Form	8
Main Desktop	6
What-What Mzansi	4
OpenSpell	7
Themba's Journey	6
Hardware (Keyboard)	2

We found the nine participants encountered a total of 39 usability and direct accessibility problems. Thirty seven of these were software problems which affected task execution by participants, while two involved the keyboard. The number of software problems affecting the various interfaces/applications ranged between four and eight (Table 4).

The number of problems affecting the login screen, the registration form and the main desktop constituted 54% of the total software problems. This corresponds with the problems identified by expert evaluators during the formal heuristic evaluation. As we discussed a proportionally high number of problems also related to the login screen and the main desktop (section 6.1.2).

Notable among the problems encountered by participants during the field evaluation were the following *usability problems*:

- The login screen did not provide any information to users who had just created user accounts or those with existing accounts regarding what should be done to access content on the DD. Login information is explicitly addressed to guest users and those wishing to create new user accounts (see Figure 3). However, after creating user accounts four participants typed 'new' in the username textbox and three typed 'dd1' (the common password for guest access),

Digital Doorway evaluation – Task list

1. Read the 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 13. Task list for the field evaluation (using the application Themba's Journey)

rather than type their user account information. This is because the majority of users typically access DD content as guests.

- The DD provided no feedback following incorrect username and/or password, but repeatedly returned the same screen. The faces of six of the eight participants, who created new user accounts, expressed confusion as they could not discern the reason for the return to the same screen.
- The registration form did not indicate any of the data fields as being mandatory or optional. Hence, five participants chose not to provide data for the home language and preferred language fields. This resulted in repeated error messages urging them to fill each field in turn.
- An error message that contradicted the hint provided next to the password field was triggered after participants typed passwords with length fewer than six characters. Although the hint urged users to choose between "6 and 10 characters", the error message was "Passwords must be between 6 and 14 characters".
- Accidental click on the <Cancel> button, located close to the <Register User> button, resulted in closure of the form, without any prompt for participants to confirm a request to cancel the registration process. Thus all the data entered by two participants up to that point was erased.
- Only three participants were able to locate the volume control button, represented by the green slider bar on the taskbar.
- The three participants whose tasks involved the use of the *What-What Mzansi*, could not access the game instructions. The instructions can be accessed by clicking the <?> icon. Participants expected to find the instructions in <About> menu.
- The default mode at the start of *What-What Mzansi* was not a full screen view. This resulted in some of the control buttons and the character that reads out the instructions being hidden from users' view. Three participants were unable to change to a full screen view without assistance.
- In *OpenSpell*, two participants could not associate the (*) symbols with the application's level of difficulty. Hence, they could not select their desired level of difficulty without assistance.

As we mentioned the DD at the school was installed in an open area (see section 6.2.1) which revealed several *direct accessibility problems*:

- The poor quality of the voice output meant participants had to lean forward to keep their ears very close to the screen to hear what was being said. This problem happened even when the volume was set at the highest level and was compounded by the sometimes noisy environment of usage.
- Participants could not use the keyboard to provide input during spelling exercise. Two participants initially tried to provide input via the keyboard, but they could only use the onscreen keyboard in *OpenSpell*.
- The reflection from the sun, due to the inadequate provision of shading from the sun, affected the visibility of pictures displayed in *OpenSpell*.
- The background of *Themba's Journey* is very dark. The problem was worsened by excessive natural lighting and constant glaring from the sun.
- Three participants were unable to access the English equivalent of texts in *Themba's Journey* without assistance. Users are expected to hover the pointer on speech bubbles to reveal the English equivalents of the

texts. The game instructions did not specify this requirement.

Some of the *cognitive access* problems (another digital divide factor) experienced during the field study includes:

- There were confusions between the two terms ‘username’ and ‘surname’. Five participants typed their surname in the username textbox even when this was not the chosen username.
- Four participants were unable to locate the educational games required for their task. This was because the applications were placed in a folder named ‘new_content’ on the desktop. Participants expected to find the applications under the ‘Game’ sub-menu.

6.2.3 Participant Questionnaire Process

We used a questionnaire to complement the other evaluation methods and compared the results obtained with that of the field observations. We derived questions from a selection of the heuristics that emerged from our literature investigation (see Appendix B). Participants used a five-point Likert scale to rate positive statements relating to the usability of the selected interfaces and applications. In addition to the rating scale, each question asked participants to make open-ended comments on problems experienced during their interaction with the DD.

We gave each of the nine field evaluation participants the questionnaire after undertaking the evaluation exercises. We had intended participants to complete and return the questionnaires immediately after evaluation sessions; however, all participants in turn requested that they take the questionnaire home and return it the following day. We conducted sessions in the afternoons, to prevent disruptions to learning activities as much as possible. If we had insisted on completing the questionnaire there and then participants would leave the school much later, which could compromise their safety. Thus, we granted participants’ request to take the questionnaires home but this reduced the return rate, and three questionnaires were never returned (see Table 3). Four questionnaires were returned the day after the evaluation, one was returned two days later, while another was only returned after several weeks.

6.2.4 Participant Questionnaires Results

We analyzed returned questionnaires by summarizing the participants’ ratings to specific statements in the questionnaires. We grouped participants’ open-ended comments, verbatim, according to similar problems. Then, in order to verify and clarify the ratings and comments, we compared participants’ ratings and comments on the questionnaires with their behaviour during the evaluation sessions.

Participants’ responses to the questionnaires were mixed. While some gave little or no response to the questionnaires, others provided valuable insight on their view of the usability of the DD, as well as the nature of problems they encountered during their interactions with the DD. Out of the six participants who returned their questionnaires, three did not provide ratings to several of the statements, and when they provided ratings, their answers always varied between *strongly agree*, *agree*, or *neither agree nor disagree*. Their positive responses contradicted the behaviour we 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 rated all of the statements and provided additional qualitative feedback. While some of their ratings were positive,

many others were negative. For example, out of the 23 statements in the questionnaire, 14 were rated with *disagree* or *strongly disagree* while nine were rated with *strongly agree*, *agree*, or *neither agree nor disagree*. The ratings by these three participants were also in line with our observations during the evaluation sessions. Some of the additional comments by these three participants are reproduced below:

- “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”.
- “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”.
- In response to a statement relating to error recognition, a participant wrote “Not at all times, sometimes you don’t even know you’ve made a mistake”.
- “No, they have to be written, not putting symbols and expect us to find where is what I mean that’s not good”. This was in response to a statement on their ability to interpret the meaning of symbols and icons on the DD.

Our field evaluation provided several important insights. Field observations revealed problems that correlated with those detected by the expert evaluators and, when compared with the self-reported problems in questionnaires show the value of careful observation of a few participants. We discuss these and other insights more deeply in the next section.

7. REFLECTION ON THE USE OF INSTRUMENT IN EVALUATING THE DIGITAL DOORWAY

To conclude our design research process we reflect on the outcome of the evaluations by comparing the results of the three evaluation methods used, comment on the effectiveness of the use of the multi-category heuristics developed, and make recommendations on the evaluation methods to be used by the DD team.

7.1 Comparing the Evaluation Methods

In this section, we reflect on the effectiveness of utilizing the heuristic evaluation method, field observations and the questionnaire to validate the developed heuristics.

7.1.1 Comparing Heuristic Evaluation and Field Observations

We compared the set of usability and direct accessibility problems detected by the expert evaluators with the problems experienced by participants in the field observations. A large number of potential usability and direct accessibility problems were identified by the team of experts (an aggregate of 71 problems), while the field observation revealed actual problems that affected participants’ interaction with the DD (39 in total). This distinction often arises with heuristic evaluation. We list some of the similarities and differences in Appendix C and, here, discuss how the involvement of users brought to light problems that were not revealed through the heuristic evaluation method.

The heuristic evaluation method revealed both predictable and, arguably, inconsequential problems. The field participants did encounter many of the problems identified by expert evaluators. For example, evaluators flagged, as problematic, the failure of the system to provide any form of feedback to users who entered an incorrect username and/or password at login. Correspondingly, six field participants were unable to determine the reason why they were returned to the same login screen

again and again. Participants repeatedly provided usernames and/or passwords that the system did not recognize and we had to intervene after several attempts to inform participants of this.

Sometimes the experts identified problems that were inconsequential to the group of participants. For example, the experts detected inconsistencies between the use of uppercase and lowercase letters for the caption of desktop elements but these did not impact negatively on users' interactions with the DD, since they related to the aesthetics of the user interface.

The field observation also uncovered problems that were not detected by the experts, bringing to light twelve additional

Table 5. Additional problems discovered during field usability evaluation

Problem	Applicable heuristics
1. 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.	1.1 1.3
2. Confusion between the terms 'surname' and 'username'. <i>Comment:</i> Primarily due to low level of computer literacy of the participants.	None
3. 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". <i>Comment:</i> The DD keyboard does not have function keys. Code was probably intended for another application. Proper testing before deployment will uncover this type of problem.	None
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.	1.1
5. Fields with input errors were not highlighted.	2.9
6. Close proximity of the <Register User> and <Cancel> buttons.	1.15 1.16
7. No message for user confirmation following unintended click on <Cancel> button before closing the form.	1.15 1.16
8. Frequent screen flicker when launching educational game applications.	1.14
9. Unusually large icons due to the configuration of one of the terminals, resulting in control buttons being hidden from viewers. <i>Comment:</i> Proper testing before deployment is required to uncover this.	None
10. The reflection from the sun affected the visibility of pictures displayed by the application. <i>Comment:</i> Evaluation at natural environment of use is required to uncover this type of problem	None
11. Very dark background, exacerbated by the open area in which the DD was installed, reduced visibility.	3.9
12. Inability to execute the 'Walk' option at one of the crossroads in the application. <i>Comment:</i> Proper testing before deployment is required to uncover this type of problem.	None

problems (Table 5). For example, the expert evaluators identified the least number of problems in the registration form (see Table 2), yet the highest number of problems encountered by participants occurred while they were using the registration form (Table 4). Six of the additional problems could have been picked up by the experts using the multi-category heuristics. However, the remaining six problems could not have been identified by a heuristic from the set of multi-category heuristics developed. As shown in Table 5 (which references the problems to the heuristic number in Appendix A) the problems related to the specific environment of use, the configuration of the particular DD terminal at the school, and the low level of computer literacy of the study participants. For example, a common problem of confusion between terms 'surname' and 'username' would be difficult to detect by experts using the heuristic evaluation method. Furthermore, the DD at the school was installed in an open area, with insufficient shading from direct sunlight. This affected the visibility of the graphics in *OpenSpell* and *Themba's Journey*, making them virtually unusable. Thus, the problems revealed by the field evaluation show the importance of involving the target user group in the actual use context.

7.1.2 Comparing Field Observations and Responses to Questionnaires

Although questionnaires are commonly used in HCI research, its sole use as an evaluation method may not give true reflection of the usability of a specific application [4, 28]. A common disadvantage of the questionnaire evaluation method is low return rate, since it relies on participants to complete and return the questionnaire. In our study only six of the nine field evaluation participants returned their questionnaires and of these, three did not rate many of the statements in the questionnaire. Another shortcoming of evaluation by questionnaire is that respondents sometimes provide answers which they deem to be socially acceptable and actual user behaviour should take precedence over responses to questionnaires [28, 35]. This was confirmed in this study where three participants provided responses that contradicted our observations of their interactions with the DD.

An advantage of observation over the questionnaire method involved immediate availability of evaluation data. The evaluation sessions were captured live, unlike the questionnaires where we relied on participants to return the completed questionnaire the day after their sessions.

The questionnaire evaluation method did not reveal any additional problems that were not picked up by observation. However, the semi-structured nature of the questionnaire gave participants the opportunity to provide qualitative comments regarding the elements they found to be problematic during the evaluation sessions. A short interview after a field observation can supply similar information.

7.1.3 Comparing Heuristic Evaluation and Responses to Questionnaires

The use of heuristic evaluation method in this study was more effective than the questionnaires in identifying potential and real usability and direct accessibility problems.

As stated in section 6.1.2, expert evaluators identified an aggregate of 71 usability and direct accessibility problems, but participants' responses to the questionnaire indicated that the usability of the DD was good in general. For example, five of the six participants who returned their questionnaires rated the icons and symbols used on the DD as being easy to interpret; while expert evaluators found them to be problematic (this was also noted during observations). The questionnaire evaluation

method did not reveal any new problem that has not been identified by the expert evaluators using the multi-category heuristics.

7.1.4 Correctness and Coverage of Heuristics

As a final step in determining the effectiveness of the application of the multi-category heuristics, we assessed them against Sim et al. [48] criteria of *correctness of terminology*, and *coverage and thoroughness*:

- *Correctness of terminology*: The multi-category heuristics were assessed by a total of six experts who had previous experience in usability and accessibility evaluation of interactive systems during development. The evaluation process did not uncover inadequacies in the terms and language used to describe the heuristics.
- *Coverage and thoroughness*: To ensure that the heuristics provide adequate coverage for the DD context, 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. During the formal evaluation, the set of multi-category heuristics helped to uncover 71 usability and direct accessibility problems in the DD applications and interfaces assessed and were found to cover all the usability problems experienced during the field observation, barring the six that were related to the context of use.

7.2 Evaluation Methods Best Suited for the Digital Doorway Team

The primary objective of our research was to develop an instrument that could guide the DD team and contract developers in the design and evaluation of applications developed for the DD. The remainder of this section presents our recommendations, based on the lessons learnt during our study, on suitable usability and direct accessibility evaluation methods that could be utilized by the DD team in the design and evaluation of applications for the DD.

Results from the three evaluation methods showed the complementary role of the various methods. However, of the three methods, evaluation by questionnaires was the least effective as the method did not yield additional value to the evaluation exercise when compared with the other two methods. Not only did we experience problems with the return rate, but it was also less effective due to discrepancies in the information provided by respondents and the observations made during actual use. Due to these reliability problems we would not recommend the use of end-user questionnaires on its own by the DD team in evaluating DD applications and interfaces. Post-field observation interviews should rather be used. Both the heuristic evaluation, based on the set of multi-category heuristics developed, and the field observations were found to be very effective in identifying possible and real usability and direct accessibility problems. The use of the multi-category heuristics is recommended as first level design and formative evaluation guidelines for the developers.

The heuristic evaluation method is considered suitable for use by both expert and non-expert evaluators [12]. In the case of non-experts, heuristics of a higher granularity is required. Hence, the granularity of our multi-category heuristics are much higher than, for example, the set of ten general heuristics proposed by Nielsen [36]. Although the DD team members have no usability expertise at present, they are experienced software developers and the detailed set of multi-category heuristics developed in our study can assist the developers in the design and a subsequent heuristic evaluation of the

applications they develop. Using the heuristics as guidelines during the development phase will also ensure that many potential usability and direct accessibility problems can be identified and corrected before end-users are exposed to the applications. If used as a summative evaluation mechanism, multiple evaluators (more than three as was explained in section 6.1.2) will be required for the heuristics to be optimally effective, and should also include expert usability and accessibility evaluators.

Evaluations by the DD team should, however, not replace evaluating the software with end-users. Results from the field evaluation with end-users have shown that critical usability and direct accessibility problems, especially those that are context related, can be overlooked in heuristic evaluations, even when conducted by experts and even though the heuristics are considered correct and adequate in coverage. We therefore recommend that summative field evaluations should also be conducted with end-users of the DD on-site. This may require contracting external expert evaluators to facilitate such evaluations.

This concluded the final phase of the design research process, resulting in the compilation of the research reports. The findings were communicated to high-level managers and DD project team members at the CSIR Meraka Institute, and were received positively.

8. CONCLUSION

The primary contribution of our research was the set of multi-category heuristics developed for evaluating non-standard systems, such as the DD. The heuristic set can be refined in future for use by developers in similar contexts as the DD team. The usefulness of the heuristics was tested by five expert evaluators, who utilized the set in a formal heuristic evaluation of the DD. We found the heuristic evaluation method, based on the set of multi-category heuristics developed, to be a useful, effective and appropriate evaluation method for the DD because of its ability to uncover large numbers of usability problems even before deployment. Also, because the DD does not support the use of assistive devices in its current configuration, which made the inclusion of people with disabilities as participants impossible, the heuristic evaluation method proved to be the most appropriate method to evaluate the direct accessibility support provided in the system.

But we also demonstrated that the outcome of single evaluation method should not be considered as comprehensive and complete. In our review of usability and accessibility evaluation methods, it was evident that each method has associated benefits and limitations. In practice, evaluation methods are typically combined to take advantage of the various methods and offset their inherent limitations [35, 43]. In our study, the results obtained through the utilization of complementary evaluation methods also proved beneficial in that it showed that although the heuristic evaluation would be an appropriate and effective formative evaluation method, it should be combined with end-user studies in order to be fully comprehensive. Through the heuristic evaluation method, based on the multi-category heuristics developed, a substantial number of potential usability and/or direct accessibility problems were identified, while the field evaluation revealed actual problems that were encountered by participants (although the number of problems identified was not as high as with the heuristic evaluation). Because the evaluation was conducted at the natural environment of use, we were also able to record specific problems related to the environment of use.

The study in its wider context also highlighted the important role of usability and accessibility in the effort to bridge the

digital divide. The digital divide is more than the acquisition of computing devices. Appropriate usability and accessibility, as well as the provision of content that is relevant to the community [10, 58] are just as important. For example, *Themba's Journey*, one of the educational applications evaluated, is provided in the IsiXhosa and English languages. However, the usability of the English equivalent is compromised considerably by a poor design decision, which requires users to hover the mouse pointer on speech bubbles in order to access the content, severely affecting the accessibility of the content to non-IsiXhosa speakers. The same can be said of other applications not offered in languages appropriate to the region in which a particular DD is deployed.

The resulting set of multi-category heuristics can act as guidelines for the DD software developers to improve the usability and direct accessibility of the applications/interfaces they develop in future, and can also be used in the first-level, formative evaluation of such software. If used in conjunction with a summative evaluation involving end-users, it would also benefit the wider user population of the DD, in the form of more user-friendly DD system.

In deciding on the appropriate evaluation methods and conducting the actual evaluation, we utilized the phases of a design research, since it mapped well to the iterative activities carried out in the study. To the wider HCI research community, our design research approach offers a novel systematic methodology for the design of application-specific heuristics.

Future research will investigate the impact of the application-specific multi-category heuristics in guiding the design decisions of DD developers in developing new software for the DD or improved versions of the current software. The aim is also to extend the heuristics to include the evaluation of the interaction devices (hardware) of the current configuration of the DD and to inform future development of such hardware. Long-term studies are also required to investigate other usability principles, such as the learnability and memorability of particular applications and interfaces. This would require current problems with uniquely identifying the DD users to be sorted out first though.

9. ACKNOWLEDGMENTS

We thank the DD team at the CSIR Meraka Institute for making the system available for the evaluation exercise. We also deeply appreciate the expert evaluators and the learners who took part in the field evaluation, as well as Dr Nicola Bidwell for her guidance in finalising the paper.

REFERENCES

- [1] Adebessin, F., Kotzé, P. and Gelderblom, H.: The Complementary Role of Two Evaluation Methods in the Usability and Accessibility Evaluation of a Non-Standard System. In: (eds.): Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists South Africa: Bela-Bela, 2010, DOI: 10.1145/1899503.1899504
- [2] Adebessin, T.F.: Usability and Accessibility Evaluation of the Digital Doorway. MSc Dissertation, School of Computing, UNISA, 2011.
- [3] Alessi, S.M. and Trollip, S.R.: *Multimedia for Learning: Methods and Development*. 3rd ed. Allyn & Bacon, Massachusetts, (2001).
- [4] Ammenwerth, E., Kaiser, F., Wilhelmy, I. and Höfer, S.: Evaluation of User Acceptance of Information Systems in Healthcare - the Value of Questionnaires. In *Medical Informatics Europe (MIE 2003)*. 2003. St. Malo, France.
- [5] Barnum, C.M.: *Usability Testing and Research*. Allyn & Bacon, Massachusetts, (2002).
- [6] Bayazit, N.: Investigating Design: A Review of Forty Years of Design Research. *Design Issues*. 20(1) (2004), 16-29, DOI: 10.1162/074793604772937379.
- [7] Bertot, J.C.: The multiple Dimensions of the Digital Divide: More than the Technology 'Haves' and 'Have Nots'. *Government Information Quarterly*. 20(2) (2003), 185-191, DOI: 10.1016/S0740-624X(03)00036-4.
- [8] Bessière, K., Newhagen, J.E., Robinson, J.P. and Shneiderman, B.: A Model for Computer Frustration: The Role of Instrumental and Dispositional Factors on Incident, Session, and Post-session Frustration and Mood. *Computers in Human Behavior*. 22 (2006), 941-961, DOI: 10.1016/j.chb.2004.03.015.
- [9] bridges.org: Digital Divide. (2006) [cited 31 May 2010]; Available from: http://www.bridges.org/digital_divide.
- [10] bridges.org: Real Access / Real Impact Criteria. (n.d) [cited 31 May 2010]; Available from: http://www.bridges.org/Real_Access.
- [11] Cambridge, G.: Digital Doorway: Enriching your Mind. In 2nd CSIR Biennial Conference. 2008.
- [12] Chattratichart, J. and Lindgaard, G.: A comparative evaluation of heuristic-based usability inspection methods. In: (eds.): CHI '08 Extended Abstracts on Human factors in Computing Systems. ACM, New York, 2008, 2213-2220, DOI: 10.1145/1358628.1358654
- [13] Cockton, G., Lavery, D. and Woolrych, A.: Inspection-Based Evaluations. In: A. Sears and J.A. Jacko, (eds.): *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*. Lawrence Erlbaum Associates, New York, 2008.
- [14] Desurvire, H., Caplan, M. and Toth, J.A.: Using Heuristics to Evaluate the Playability of Games. In: (eds.): Conference on Human Factors in Computing Systems, CHI 2004. ACM, Vienna: Austria, 2004, 1509-1512, DOI: 10.1145/985921.986102
- [15] Deurvire, H. and Thomas, J.C.: Enhancing the Performance of Interface Evaluators Using Non-Empirical Usability Methods. In: (eds.): 37th Annual Meeting of the Human Factors and Ergonomics Society. 1993.
- [16] Dix, A., Finlay, J., Abowd, G.D. and Beale, R.: *Human-Computer Interaction*. 3rd ed. Pearson Education Ltd, London, (2004).
- [17] Dumas, J.S.: User-Based Evaluation. In: J.A. Jacko and A. Sears, (eds.): *The Human-Computer Interaction Handbook*. Lawrence Erlbaum Associates, Mahwah, 2003.
- [18] Dumas, J.S. and Fox, J.E.: Usability Testing: Current Practice and Future Directions. In: A. Sears and J.A. Jacko, (eds.): *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*. Lawrence Erlbaum Associates, New York, 2008, DOI: 10.1201/9781410615862.ch57.
- [19] Gelderblom, J.H.: Designing Technology for Young Children: Guidelines Grounded in a Literature Investigation on Child Development and Children's Technology. PhD Thesis, School of Computing, UNISA, 2008.
- [20] Greeff, M. and Kotzé, P.: A Lightweight Methodology to Improve Web Accessibility. In: (eds.): Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists. South Africa: Vanderbijlpark, 2009, 30-39, DOI: 10.1145/1632149.1632155
- [21] Greyling, E. and Smith, R.: An innovative ICT Solution to Steer Rural Communities to Global Understanding: A Case

- Study From Durban, South Africa. In: (eds.): World Library and Information Congress: 74th IFLA General Conference and Council. Canada: Quebec, 2008, 1-12.
- [22] Gush, K., Cambridge, G. and Smith, R.: The Digital Doorway - Minimally Invasive Education in Africa. In: (eds.): ICT in Education Conference. Cape Town, 2004, 1-7.
- [23] Gush, K. and De Villiers, M.R.: Application Usage of Unsupervised Digital Doorway Computer Kiosks in Remote Locations in South Africa. In: (eds.): Annual Research Conference of the South African Institute of Computer Scientists and Information Technologist. South Africa: Bela-Bela, 2010, 93-103, DOI: 10.1145/1899503.1899514
- [24] Gush, K., De Villiers, M.R., Smith, R. and Cambridge, G.: Digital Doorways. In: J. Steyn, J.P. Belle and E.V. Mansilla, (eds.): ICTs for Global Development and Sustainability: Practice and Applications. IGI Global, Pennsylvania, 2010, 96-126, DOI: 10.4018/978-1-61520-997-2.ch005.
- [25] Henry, S.L.: Understanding Web Accessibility. In: J. Thatcher, et al., (eds.): Constructing Accessibility Web Sites. Glasshaus, Birmingham, 2002.
- [26] Henry, S.L.: Just Ask: Integrating Accessibility Throughout Design. 2007.
- [27] Hollingsed, T. and Novick, D.G.: Usability Inspection Methods after 15 Years of Research and Practice. In: (eds.): The 25th Annual ACM International Conference of Communication. ACM, 2007, DOI: 10.1145/1297144.1297200
- [28] Holzinger, A.: Usability Engineering Methods for Software Developers. *Communications of the ACM*. 48(1) (2005), 71-74, DOI: 10.1145/1039539.1039541
- [29] IBM: IBM Software Accessibility Checklist. (2009) [cited 24 November 2009]; Available from: <http://alturl.com/9xw2>.
- [30] International Organization for Standardization ISO 9241-11 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 11: Guidance on Usability. 1998.
- [31] Malone, T.W.: What Makes Things Fun to Learn? Heuristics for Designing Instructional Computer Games. In 3rd ACM SIGSMALL Symposium and 1st SIGPC Symposium on Small Systems. 1980. USA: California, DOI: 10.1145/800088.802839
- [32] Malone, T.W.: Toward a Theory of Intrinsically Motivating Instruction. *Cognitive Science*. 5(4) (1981), 333-369.
- [33] Mayhew, D.J.: Principles and Guidelines in Software User Interface Design. Prentice Hall, Inc, New Jersey, (1992).
- [34] Meraka Institute: The Digital Doorway: Enriching Your Mind. (n.d) [cited 10 July 2010]; Available from: <http://alturl.com/ktun5>.
- [35] Nielsen, J.: Usability Engineering. Academic Press, Inc, California, (1993).
- [36] Nielsen, J.: Heuristic Evaluation. In: J. Nielsen and R.L. Mack, (eds.): Usability Inspection Methods. John Wiley & Sons, Inc., New York, 1994.
- [37] Nielsen, J.: Usability 101: Introduction to Usability. (2003) [cited 20 July 2009]; Available from: <http://www.useit.com/alertbox/20030825.html>.
- [38] Nielsen, J.: Digital Divide: The Three Stages. (2006) [cited 30 April 2010]; Available from: <http://www.useit.com/alertbox/digital-divide.html>.
- [39] Norman, D.A.: The Design of Everyday Things. MIT Press, London, (2001).
- [40] Ozok, A.A.: Survey Design and Implementation in HCI. In: A. Sears and J.A. Jacko, (eds.): The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications. Lawrence Erlbaum Associates, New York, 2008.
- [41] Pierotti, D.: Heuristic Evaluation - A System Checklist. (n.d) [cited 11 November 2011]; Available from: <http://alturl.com/cq283>.
- [42] Pretorius, M.C., Calitz, A.P. and Van Greunen, D.: The Added Value of Eye Tracking in the Usability Evaluation of a Network Management Tool. In: (eds.): Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT 2005). 2005, 1-10.
- [43] Rogers, Y., Sharp, H. and Preece, J.: Interaction Design: Beyond Human-Computer Interaction. 2nd ed. John Wiley & Sons Ltd, Chichester, (2007).
- [44] Rogers, Y., Sharp, H. and Preece, J.: Interaction Design: Beyond Human-Computer Interaction. 3rd ed. John Wiley & Sons Ltd, Chichester, (2011).
- [45] Rubin, J. and Chisnell, D.: Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. 2nd ed. Wiley Publishing Inc., Indiana, (2008).
- [46] Shelley, B.: Guidelines for Developing Successful Games. (2001) [cited 10 March 2010]; Available from: <http://alturl.com/rcg5r>.
- [47] Shneiderman, B.: Designing the User Interface: strategies for Effective Human-Computer Interaction. 3rd ed. Addison Wesley Longman, Inc., California, (1998).
- [48] Sim, G., Read, J.C. and Cockton, G.: Evidence Based Design of Heuristics for Computer Assisted Assessment In: T. Gross, et al., (eds.): Human-Computer Interaction - INTERACT 2009. SpringerLink, 2009, DOI: 10.1007/978-3-642-03655-2_25
- [49] Smith, R., Cambridge, G. and Gush, K.: Digital Doorway Computer Literacy through Unassisted Learning in South Africa. In 1st CSIR Biennial Conference. 2006. CSIR Convention Center Pretoria.
- [50] Smith, S.L. and Mosier, J.N.: Guidelines for Designing User Interface Software. (1986) [cited 14 November 2011]; Available from: <http://www.hcibib.org/sam/index.html>.
- [51] Story, M.F., Mueller, J.L. and Mace, R.L.: The Universal Design File: Designing for People of All Ages and Abilities. (1998) [cited 17 November 2009]; Available from: <http://www.eric.ed.gov/pdfs/ed460554.pdf>
- [52] Total Validator: Total Validator. (2009) [cited 08 December 2009]; Available from: <http://www.totalvalidator.com/>.
- [53] Travis, D.: 247 Web Usability Guidelines. (2009) [cited 10 November 2011]; Available from: <http://www.userfocus.co.uk/resources/guidelines.html>.
- [54] United States Access Board: Electronic and Information Technology Accessibility Standards (Section 508). (2000) [cited 09 November 2009]; Available from: <http://www.access-board.gov/sec508/standards.htm>.
- [55] Vaishnavi, V. and Kuechler, W.: Design Research in Information Systems. (2004) [cited 17 November 2009]; Available from: <http://alturl.com/prscq>.
- [56] Walls, J.G., Wildmeyer, G.R. and El Sawy, O.A.: Building an Information System Design Theory for Vigilant EIS. *Information Systems Research*. 3(1) (1992), 36-59, DOI: 10.1287/isre.3.1.36.
- [57] Wharton, C., Braffort, J., Jeffries, R. and Franzke, M.: Applying Cognitive Walkthroughs to More Complex User Interfaces: Experiences, Issues, and Recommendations. In

SIGCHI Conference on Human factors in Computing Systems 1992, DOI: 10.1145/142750.142864
 [58] Wilson, E.J.: The Information Revolution and Developing Countries. MIT Press, London, (2006).

[59] World Wide Web Consortium: Web Content Accessibility Guidelines 1.0. (1999) [cited 7 May 2009]; Available from: <http://www.w3.org/tr/wai-webcontent/>.

Appendix A. Heuristics set for evaluating the Digital Doorway

Category 1: General Usability Heuristics	
1.1	Provide information that will enable users understand how to interact with the Digital Doorway using clear and simple terminology.
1.2	Provide clear indication of what the next required action is.
1.3	Avoid the use of technical terms.
1.4	Icons, symbols and menu items should be labelled with intuitive and meaningful names, taking into account user context and experience.
1.5	Ensure that information sequence and layout appear in natural and logical order.
1.6	Follow and adhere to platform and industry standards and conventions.
1.7	The mappings between controls and their effect should be intuitive and easily understood.
1.8	Users should not use considerable cognitive resources trying to interpret the meaning of icons, menus and symbols, and to navigate the interface.
1.9	Objects, options and permissible actions should be visible so that users do not have to remember instructions.
1.10	Users should not be required to remember information from one screen to another.
1.11	Menu options should be logically grouped to aid the recognition of available functionalities.
1.12	Feedback should be provided in clear and unambiguous terms.
1.13	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.
1.14	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.
1.15	Provide support for system exploration by the user by allowing easy reversal of actions.
1.16	Prevent user error by using appropriate constraints at strategic points.
1.17	Error messages should be context-specific in relation to the action performed.
1.18	Error messages should be given in language comprehensible to users, not using technical terms.
1.19	Error messages should precisely describe what the problem is and offer ways of solving them.
1.20	Ensure that the interface is not cluttered with irrelevant information, control buttons and icons.
1.21	Provide information and control options close to when the user is required to make use of them.
1.22	Be consistent in the naming conventions used for icons, symbols and objects.
1.23	Make sure that the same terms, actions or symbols mean the same thing across applications.
1.24	Create the same 'look and feel' effect across applications so users can extend knowledge to similar situations.
1.25	Ensure that labels/titles for icons, menus and symbols accurately describe their content.
1.26	Component labels/titles should not mislead users into accessing content they would otherwise not be interested in.
1.27	The Digital Doorway should not impose unnecessary constraints on the user input method.
1.28	Where user input can be provided via the keyboard and onscreen keys, the user should be allowed to provide input through either method.
1.29	The Digital Doorway should support multiple output methods.
Category 2: Form Usability Heuristics	
2.1	Provide visible cue by positioning the cursor in the first field at start of the form.
2.2	Cursor movement should follow the order in which form elements are organized.
2.3	Users should be able to edit data fields by moving the cursor backward and forward, rather than having to retype the whole field.
2.4	Ensure that related items are grouped together to aid readability.

2.5	Provide visual reinforcement for element groups through efficient use of white spaces and borders.
2.6	Ensure that required information is clearly specified.
2.7	Designate required fields in standard and consistent ways, taking into account the user's age and knowledge.
2.8	Give feedback for missing data fields in clear and unambiguous terms, taking into account the user's age and knowledge.
2.9	When input errors are detected, the cursor should be positioned in the error field with the field highlighted to attract the user's attention.
2.10	Give meaningful names to field captions/labels, taking into account the user's age and experience.
2.11	Ensure that captions/labels are distinct from data entry fields.
2.12	Clearly specify the limit for data that has minimum or maximum allowable length.
Category 3: Heuristics to Support Direct Accessibility	
3.1	Font size of instructions should be large enough to enable easy perception by users with low vision.
3.2	Information should be accessible without undue physical efforts.
3.3	Important information should be clearly distinguishable from other peripheral contents.
3.4	Provide audio equivalent of instructions and information to afford access by users who cannot read.
3.5	Provide quality speech output that enable users to hear and comprehend their meanings.
3.6	Provide feedback using multiple modes to facilitate access and comprehension.
3.7	Text equivalent of graphic or audio information should convey the same message.
3.8	Ensure that colour alone is not used to represent important information.
3.9	Ensure that background and text colours contrast well with each other.
3.10	Allow keyboard navigation for operations/tasks that do not essentially require use of the mouse.
3.11	Ensure that menus and buttons are accessible using the keyboard.
3.12	Avoid automatic progression from one screen to the next for audio-visual information. Users should explicitly select forward/backward progression.
3.13	Provide controls that enable users to pause, continue, or repeat audio-visual information.
3.14	Users should be able to adjust the volume of audio information
3.15	Equivalent audio information should be synchronized with the text alternatives.
Category 4: Educational Game Usability Heuristics	
4.1	Games should have clear goals and objectives.
4.2	Ensure that learners can easily determine whether they are getting closer to the goal.
4.3	Provide an easily accessible instruction on how to play the game.
4.4	Permissible actions should be clearly specified.
4.5	Clearly specify constraints and restrictions governing the game.
4.6	Learners should be able to adjust the game's level of difficulty.
4.7	The application should be able to adjust the level of difficulty based on the learner's performance.
4.8	Whenever appropriate, give learners the option of returning to where they left off when the program is temporarily exited.
4.9	Ensure constant challenge through adjustable difficulty level.
4.10	Performance feedback should not be given using negative or sarcastic statements.
4.11	Provide constructive and corrective feedback that will enable the learner to learn from mistakes and improve future performance.
4.12	Learners should have more than one opportunity to provide answers.
4.13	Appropriate hints should be provided for the correction of cognitive mistakes.
4.14	All control mechanisms should be visible and easily accessible.
4.15	Provide learner control options for forward progression to facilitate skipping a section and backward progression, which enables the review of a previous section.
4.16	All control mechanisms should be easy to use without requiring undue physical efforts.
4.17	Provide clear exit mechanism to allow learners leave the game at any stage.
4.18	Request to terminate the program should be confirmed by the learner to avoid unintentional exit.
4.19	Game should be accessible in different languages.

4.20	Game content should not be biased against specific cultural or gender groups.
4.21	Game activities should be embedded in scenarios that learners can relate to.

Appendix B: User-administered questionnaire

General ease of using the Digital Doorway	
1	Instructions about how to use the Digital Doorway are clear to me.
2	Instructions and information are given in various ways like written words, spoken words, and through pictures.
3	I can change the volume of spoken words and instructions to suit me.
4	Whenever I make a mistake, the Digital Doorway tells me exactly what the mistake is in a way that I can understand.
5	The Digital Doorway clearly shows me what I need to do next so that I can complete what I am doing.
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.
7	I understand the meaning of icons (pictures) used in the Digital Doorway.
8	I am able to determine the meaning and purpose of signs and symbols used in the Digital Doorway.
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).
10	The Digital Doorway contains words used by computer people which I do not fully understand.
11	The Digital Doorway shows me how to correct my mistake.
12	The Digital Doorway gives response within a short time.
13	Spoken instructions and information are loud and clear enough for me to hear and understand.
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.
15	The space provided is sufficient for the information I need to write.
16	I am able to determine where I should write information in the form.
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.
18	It is easy for me to make corrections to information that I have written earlier without me having to retype all over again.
Usefulness of educational games in the Digital Doorway	
19	The Digital Doorway does not make fun of me when my answer is wrong.
20	It is easy for me to choose in which language I want to play a game.
21	The Digital Doorway informs what I should do in order to play the games.
22	I am able to control how easy or difficult I want the game to be.
23	It is easy for me to determine what the computer games in the Digital Doorway will be teaching me.

Appendix C. Similarities and differences between problems identified by expert evaluators and actual user problems

Problems identified by expert evaluators	Actual user problems
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 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. Seven participants actually 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.
On the login screen, no feedback was provided when an incorrect username and/or password is entered.	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 we had to inform participants that the problem was with the username or password being provided.
Expert evaluators did not identify this as a potential problem. The problem occurred due to the low level of computer literacy among the field participants. Heuristics have limitations in revealing this type of unpredictable problem.	Five participants confused their surname with a 'username' and typed their surname in the username field, even when this was not the chosen username.
The font sizes of the instructions on how to log in or create user account in four other languages (IsiXhosa, Afrikaans, Sotho and Venda) were too small.	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 were unable to read English.
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 key without the desired effect they then pressed the other.
When the registration form was activated, the cursor was 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.
There is no indication of which fields are compulsory to be filled and which ones are optional.	Five participants left the home language and preferred language fields empty only to have error messages urging them to fill the fields.
Users 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. Using the pointing device, they positioned the insertion point over relevant fields before typing or clicking to make their selections. This was not surprising since they are not expert users and would not know of the use of the Tab key for such actions.
None of the expert evaluators flagged the close proximity of the <Register user> and <Cancel> buttons as potential usability problems.	Two participants accidentally clicked on the <Cancel> button while filling the registration form. This inadvertent error resulted in the form being closed without any warning to the user thereby erasing all the data fields input thus far.
None of the expert evaluators identified the absence of highlights on error fields as problematic.	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 he clicked the <Register user> button. Without the user realising this, he pressed the backspace button several times and erased the wrong field unintentionally.
The functionality of the volume control slider on the desktop is not clear from its look.	Only three participants were able to locate the volume control button on the desktop, the others required assistance after several failed attempts.
One folder on the desktop has the caption 'new_content'. This is not descriptive of the applications it contained. The game applications <i>What-What Mzansi</i> , <i>Themba's journey</i> and <i>OpenSpell</i> are hidden inside the folder 'new_content'.	Three participants unsuccessfully searched for the applications within the <Game> submenu, located in the Resource menu, before they were told where to find applications.
The level of contrast between the dark blue background and the grey foreground used to label icons on the desktop is low.	Three of the participants found the background colour to be too dark. On several occasions, they had to shield their faces and the

The contrast between the word ‘Digital Doorway’ and the dark blue background is poor.	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.
The caption of the desktop element labelled ‘Bluetooth_saver’ has the first character capitalized while the captions for all other icons and folders are in lowercase.	This was a minor interface aesthetic problem, which did not affect the interaction of the participants with the DD.
The locations of the following icons on the taskbar are too close to one another: the right-pointing arrow ⇒, <System> <Volume control>, and Volume control slider. Users could easily click on the right-pointing arrow ⇒ while trying to use the volume control slider, thereby closing the system unintentionally.	A user accidentally clicked on the ⇒ arrow used to exit the system while trying to locate the volume control button and the system was shut down without any warning.
At the start of <i>What-What Mzansi</i> , some of the control buttons and the character that reads out instructions and questions were hidden from users’ view. A full screen mode is activated by clicking on an icon which does not indicate this function.	Three participants were unable to switch to full screen view of the application without assistance.
This problem was not encountered by the experts as it related to the configuration of a specific terminal at the school.	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 <i>What-What Mzansi</i> , in one particular instance, the right pointing arrow ‘>’ used for forward progression was hidden from users’ view. This made it impossible for participants to repeat the level which they had just completed as was required following poor performance.
The performance feedback “don’t make me laugh” after a poor performance is cheeky and not encouraging. Some users might find it offensive.	Participants who used this application were indifferent to the performance feedback.
The use of the labels <Say>, <Guess>, and <Spell> in <i>OpenSpell</i> are not descriptive of their functionalities.	The 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.
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, two participants first attempted to use the keyboard to provide their input, only to realize later that they can only use the onscreen keyboard.
The * symbols used to represent the level of difficulty are not intuitive.	Two participants were unable to set the level of difficulty in <i>OpenSpell</i> .
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.	Three participants did not know how to get the English equivalent of the narration until they were told. Much effort was required by participants to move pointer around the speech bubbles in order to read English versions. Participants’ fingers became damp on several occasions and they had to dry their hands.
The narration voice in <i>Themba’s Journey</i> is only in IsiXhosa. Non-IsiXhosa users who cannot read will not be able to use the application.	Participants were school children and were able to read the English equivalent of the narrations.
None of the expert evaluator identified the dark background of <i>Themba’s Journey</i> as potential problem. This was because the evaluation was conducted in a close-up laboratory without the interference of sunshine glare.	Application background was very dark. 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 from the sun.