

Application Usage of Unsupervised Digital Doorway Computer Kiosks in Remote Locations in South Africa

Kim Gush

CSIR Meraka and School of Computing, UNISA
P O Box 395, Pretoria, 0001
+27 12 8413028
kgush@csir.co.za

Ruth de Villiers

School of Computing
UNISA
P O Box 392, Pretoria, 0003
+27 12 429 6559
dvillmr@unisa.ac.za

ABSTRACT

Digital Doorways are rugged computer terminals that offer unassisted learning and peer-assisted learning of basic computer skills, as well as a range of computing activities from entertainment, through education, to independent research. Sites are located in impoverished areas of South Africa at schools, colleges and public community facilities. Usage is free of charge and available to the entire community. This paper poses research questions relating to application usage data and how it relates to user demographics, in order to better understand both the user base and the nature and extent of interaction with a selected set of terminals. This study thus addresses significant issues with respect to ICT for Education and Development in the Digital Doorway context. Analysis of the data indicates notable trends and patterns, and raises certain concerns.

Categories and Subject Descriptors

K.3.0 [Computers and Education]: General – *research model*.
K.3.1 [Computers and Education]: Computer Uses in Education – *design and development, evaluation/methodology, computer-assisted learning, computer-assisted training*.
K.4 [Computers and Society].
K.6 [Management of Computing and Information Systems]

General Terms

Human Factors, Performance, Verification

Keywords

Computer literacy, Community informatics, Digital Doorway, ICT in education, ICT for development, Unassisted learning

1. INTRODUCTION

This paper highlights findings from a study aimed at obtaining an overview of the nature and extent of usage of software applications on publicly accessible, rugged 'Digital Doorway' computer systems installed in remote South African communities. The study focuses on ten sites, and involves the quantitative analysis of log files of software usage at the Digital Doorway (DD) terminals. These terminals and their embedded

software have been designed to enhance basic computer skills of users.

Throughout South Africa, and indeed the whole of Africa, increased access to computer infrastructure is needed, among others, in educational contexts. Table 1 shows the extent of computer centres in South African schools, indicating extremely low figures for the Eastern Cape, KwaZulu Natal, Limpopo and Mpumalanga.

Table 1: School computer centres in South Africa, 2009. [1]

Province	Number of Schools	With Computer Centre	% With Computer Centre
Eastern Cape	5715	596	10
Free State	1643	353	21
Gauteng	1994	1510	76
KwaZulu Natal	5835	982	17
Limpopo	3918	428	11
Mpumalanga	1540	254	16
North West	1740	391	22
Northern Cape	609	314	52
Western Cape	1466	886	60
Total	24460	5714	23

Increased access to computer technology and skills development may be accomplished in various ways including the accelerated roll out of computer laboratories and information kiosks. However, a previously little-explored alternative, which may co-exist with other support facilities, is the provision of resources for independent 'unassisted learning' or 'minimally invasive education' (MIE). This is the ethos embodied in the Digital Doorway (DD) concept. According to Dangwal [2:295], 'MIE demonstrates a special case of the interplay of information technology (computers) and learning processes and emphasises the role of self-directed and participatory learning'.

Providing infrastructure to encourage self-directed learning is all very well, however, as Amiel mentions, 'Technology is much more than hardware. It is a process that involves the complex interactions of human, social and cultural factors as well as the technical aspects' [3:31]. Key to understanding these interactions is obtaining an accurate picture of how the systems are being used. It is this usage, specifically in relation to age, gender and site context that forms the subject of the study presented in this paper.

A brief overview of related work is given in Section 2, followed by a background to the Digital Doorway project in Section 3. Section 4 outlines the research methodology, including four

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SAICSIT '10, October 11–13, 2010, Bela Bela, South Africa.
Copyright 2010 ACM 978-1-60558-950-3/10/10...\$10.00.

research questions, while Section 5 presents some of the findings from processed data, along with discussion based on research questions one, two and three. Section 6 answers the final research question, presents conclusions and notes lessons learned for future efforts.

2. RELATED WORK

In 1999, Dr Sugata Mitra of the National Institute of Information Technologies in India (NIIT) demonstrated via his innovative Hole-In-The-Wall experiment (HITW) that children were able to acquire basic computer skills through self discovery, without the assistance of a teacher [4]. From its conception, the underlying premise of the Hole-in-the-Wall installations was to offer a complementary learning tool to remote and almost forgotten locations – places removed physically from the cities or from the focus of the government’s development agenda [5].

Dr Mitra coined the term minimally invasive education to describe the learning paradigm at the HITW kiosks. Various academic studies have been conducted at these sites ([5], [6], [7]), and data from HITW sites in India and the Digital Doorway sites in South Africa is extensive ([8], [9]). Two notable Indian studies that shed more light on unassisted learning and application usage are those of deBoer [5] and Dangwal [10].

Dangwal reports on a comparative study of computer literacy, intellectual maturity and school performance before and after exposure to the HITW kiosks, the former being measured using an Icon Association Inventory (IAI) test. Computer literacy performance of the experimental group (boys and girls aged 8-14 years old) was seen to improve significantly at all testing zones, while academic performance was seen to improve in two of the four zones.

During a different study involving multiple Indian HITW sites [5], using questionnaires and interviews, some behavioural patterns emerged. The largest category of users was aged between 11 and 12 years old, and there were more boys than girls at the kiosks (girl-boy ration of 0.83). The average age of girls was younger than that of boys (11.37 vs 12.37) and it was noted that as the girls matured, they became reluctant to visit the kiosks where older boys were present.

Following the development and implementation of an icon test by colleagues of Mitra in which children were given icons to identify, so as to establish the degree to which computer learning had taken place [11], Mitra proposed a tri-part pedagogy for learners to acquire the necessary competence, with two of these bands not requiring the direct intervention of a knowledgeable teacher, and achievable through technology access and peer learning alone. This minimally invasive education, self-assisted learning and peer-assisted learning approach forms the core educational method on which both the HITW and the Digital Doorway projects are based.

A recent study looked at means of improving the collection of demographic data of DD users, as well as increasing a user’s sense of ownership of the computer [9]. The work covered in the present study gives a new perspective, in that it focuses on **usage** and quantitative analysis of how that usage relates to both user and site demographics.

3. THE DIGITAL DOORWAY PROJECT

Since 2002 the CSIR’s Meraka Institute has been involved in the Digital Doorway project, which aims to increase computer literacy of users in impoverished and remote communities of South Africa and further afield (including Lesotho and Uganda). Each of the over 200 DD sites consists of a 3-terminal or 4-terminal rugged computer kiosk housing a computer server, a

number of client machines, robust plexiglass-covered LCD screens and durable metal keyboards. The terminals, an example of which is shown in Figure 1, are bolted to the floor of public community centres or schools and encourage the development of computer skills through unassisted or peer-assisted learning without the presence of a teacher or instructor. The potential of this type of learning in the African context is discussed in [12] and [13]. Figure 1 demonstrates a 4-terminal setup, and highlights the typical clustered user interaction around the device.

Learning takes place through a process of exploration and discovery. Only free and open source content is available on the computers. The embedded software applications span a broad range of subject matter including educational games, instructional programs, reference documents, audio- and video content and basic computing tools such as a word processor and spreadsheet. The goal is basic computer literacy and narrowing the digital divide, or enabling a ‘digital transition’ [14]. The target users are both adults and children.

Observation (directly and through analysis of video footage) has been conducted on the manner in which users at the sites interact with the computers and with each other. Learning takes place both individually as users explore the system and collectively as more knowledgeable users show others what they have learned

4. RESEARCH METHODOLOGY

4.1 Aim and research questions



Figure 1: Users at a Digital Doorway

A vital component in evaluating the success or inadequacies of the DD Project as a whole, and individual sites in particular, is an adequate understanding of both user demographics and application usage. Both qualitative and quantitative research approaches may be used, with this study emphasizing the latter.

Over the years, large quantities of data on user demographics and application usage have been recorded by various means, including log files and user-generated information. A huge volume of information is available from the 200+ sites, most of it formally extracted since 2007 or 2008, although many sites have been in operation considerably longer. In addition to the processing of thousands of lines of log files, a further complexity is presented by the varying types and structures of data stored in these logs.

In analysing aspects of a terminal, the following set of dimensions is the most pertinent:

- The type of location (e.g. a library);
- The geographical situation (site);
- The categories of applications installed;

- The range of end-users in terms of age and gender.

The overall aim of the present study is to gain a representative snapshot of the nature and extent of usage of the software applications provided with DD terminals and to investigate how usage relates to user demographics.

To this end, a small, but stereotypical set of locations and sites were selected for in-depth investigation. Using purposive sampling [15], the primary researcher hand-picked ten specific sites that represented the range of DD environments and that were likely to generate valuable data (Table 2). Criteria used to select a representative set of ten sites from the approximately 210, were:

- Site diversity in terms of type of venue (e.g. school)
- Numbers of registered users per site – sites with the highest numbers
- Hit counts per user – highest hit counts
- Region – both rural and township
- Geographical – multiple provinces.

Data-processing tools have been employed or newly built to filter, categorise, combine and analyse data – see Section 4.3.

The findings of this research indicate usage patterns that will help in customising content for most effective use of scarce computer resources in needy areas. The analysis of complete sets of logged data from multiple sites and multiple end-users affords opportunities to identify aggregated trends and occurrences. The study focuses on usage of the embedded applications, not on usability of the DD, which is synergistically being investigated by another team of Meraka researchers. Usability tests and evaluations are focused on specified scopes, participants, locations and durations [7], while this research addresses the utility and usage levels of the subject matter. In addition to Hilbert's [16:1] question of *'Where should scarce design, implementation, testing, and usability evaluation resources be focused in order to produce the greatest benefit for users?' one can also enquire: 'Where should DD content resources be focused to produce the greatest benefit for users?'*

The study sets out to quantitatively answer four research questions regarding application usage data and how it relates to user demographics. The purpose is to better understand both the user base and interaction with the systems:

1. What categories of applications are used by various age groups on Digital Doorways? (See Section)
2. Does gender have an impact on extent of use and on application usage? (See Section)
3. Which applications are under-utilised by the intended target groups? (See Section)
4. What lessons are learned for future development, selection and presentation of applications? (See Section 6.1)

The study also aims to draw out further research questions that may subsequently be answered through qualitative research.

4.2 Data Collection

While there is undeniable value in the analysis of user interactions through surveys via interviews and questionnaires, it is important to know which applications users are actually accessing, not just what they say they are accessing. Furthermore, data collection via surveys is restricted to samples of participants rather than full populations. The primary data collection mechanism for this study was the extraction of data from comprehensive log files.

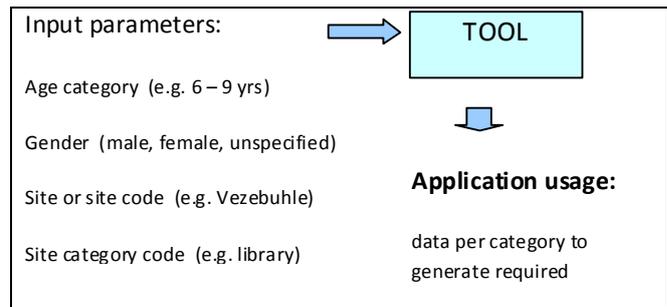


Figure 2: Log file interrogation tool (simple representation)

Each system has mechanisms for the collection and transfer of log files of user details and application usage statistics. User demographic information is collected through a user 'self-registration' procedure, whereby a user can create a unique user name and enter details of name, age, gender, home language and preferred language. Users are also able to log in as 'guest users', the disadvantage being that files saved under guests are erased daily. However, all application logs, including hits by both guest- and registered users, are stored for later retrieval. Application log files for guest users are many times larger than those for registered users, however unlike the latter, guest files are not linked to age and gender data.

Each application on the Digital Doorway is launched from a wrapper script that records the time of launch and user name of the person launching it. This information is stored in a text file on the server. All log files are compressed and transmitted daily via GPRS (cellular) backbone to a central server. This collection of compressed log files from all sites can be analysed to provide insight into application usage, especially with regard to general trends and preferences. Though studies have been undertaken on isolated observations at various sites, no previous study has been conducted on the broader scope of this present research.

While the application logs reflect the actual usage on the machines with a high degree of accuracy, the demographic data as supplied by users on self-registration is more open to error and to the possibility of entry of incorrect or incomplete information. Nevertheless the newly created software logging systems facilitate broader and better understanding of application usage (at least in terms of trends and patterns), than surveys. A similar study was done by Veeraraghavan et al. on kiosks in India [6].

4.3 Data Processing and Analysis

In the absence of a single database or data warehouse, and in order to overcome the limitations resulting from large volumes of heterogeneous data in multiple formats, new data collection and processing methods were created to analyse the raw data as the first part of this study. Data are categorised, combined and analysed using various visualisation aids. Key to the analysis of application usage data from the sites was the design of a log file interrogation tool for customisable category-specific extraction of data. A simple representation of the tool is shown in Figure 2, indicating how selected key dimensions (introduced in Section 4.1) are used as input to generate the required output. Figure 3 illustrates the main steps in moving from copious amounts of raw data to meaningful visualisations.

Manual data cleaning – the 'detecting and removing errors and inconsistencies from data in order to improve the quality of data...' [17] – was necessary to deal with limitations of the data collection mechanism, whereby users could enter certain data in

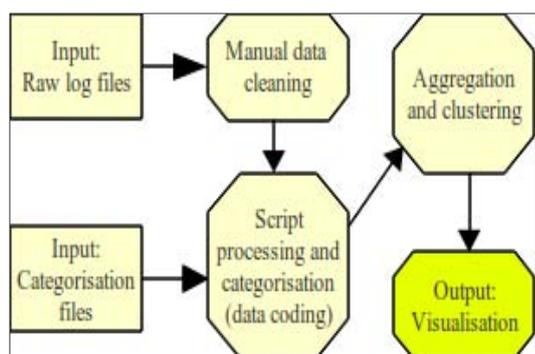


Figure 3: Process from raw data to visualization

an unrestricted way (e.g. specify age as '20' or 'twenty'). Also important to the process was manual viewing to inspect the data throughout the processing stages to verify that scripts were handling as expected and to identify areas of interest not necessarily obvious after aggregation or categorisation (e.g. an individual application having prominence of usage).

4.4 Categorisation of data

Of the approximately 210 sites installed around South Africa, ten were carefully selected for this study, based on the five criteria listed in Section 4.1. Users interacted with the DD at these ten different geographical sites and the sites, in turn, are categorised under seven types of location: school, library, etc. (Table 2).

Table 2: The ten sites in seven groupings chosen for this study

	Location	Site Name	Prov.	Township or Rural	Site Category	Tot. Reg. Users (m, f, u)	Guest User Hits	Reg. User Hits	Data collection period	Months
1	School	Kwam_Hlonipa	LIM	Rural	Education	69 (35, 12, 22)	9133	455	Apr2008–Apr2009	12
		Vezebuhle	MP	Rural	Education	690 (382, 90, 218)	16081	8755	Mar2007–Sep2007	6
2	Multi-purpose community center (MPCC)	Elandskraal	LIM	Rural	Public	258 (150, 50, 58)	17557	2758	Feb2007–Jun2008	16
3	Library	Emjindini	MP	Township	Public	474 (250, 118, 106)	8683	12525	Feb2007–Jun2009	28
		Kanyamazane	MP	Township	Public	46 (25, 6, 15)	5319	421	Feb2008–Jun2008	4
4	FET College	Letaba FET College 2	LIM	Township	Education	66 (34, 9, 23)	14798	951	Jan2007–Feb2009	13
		Letaba FET Giyani	LIM	Township	Education	306 (187, 50, 69)	3765	7598	Jan2007–Sep2007	8
5	Cust. Care Centre	Msunduzi	KZN	Rural	Public	101 (38, 17, 46)	17752	1841	Feb2007–Sep2008	19
6	Fablab	Soshunguve	GAU	Township	Public	100 (64, 12, 24)	37074	4628	Jan2007–Mar2008	14
7	Informal Market	Kagung	NC	Rural	Public	40 (25, 1, 14)	4253	1899	Jul2008–Mav2009	10
excludes names registered multiple times, m=male, f=female, u=unspecified LIM = Limpopo, MP = Mpumalanga, KZN = KwazuluNatal, GAU = Gauteng, NC = Northern Cape						2150 (1190, 365, 595)				

Provinces and the township or rural location of sites are also shown. Unless otherwise indicated, results shown in the tables and figures are based on these ten sites. Analysis of these ten sites forms part of a more extensive study involving both qualitative analysis and quantitative analysis of these and certain

other sites. To make the large quantities of data manageable, applications and ages were also grouped into categories.

The 992 applications and content sources available to users were grouped into seven general categories (Table 3) as well as 26 specific categories (see Figure 9 later). User ages (as recorded during self-registration) were grouped into seven categories (see Table 4). In total there were over **2000** users and over **30,000** application launches (hit-counts) distributed over the sites in the two years for which data was logged. The categorisation step is itself a form of interpretation [18] and, for this reason, the demographic figures and tables are in Section 5, Results and Analysis. A balance had to be struck between grouping applications into just a few broad categories – obscuring detail – and using multiple smaller categories – adding complexity to and reducing generalizability of the analysis.

Table 3: Broad application categories

Category	Description
Edutainment	Games with educational elements incorporated
Education	Software created purely for educational purposes
System	Includes file navigation and DD homepage
Office	Office suite including word processor, spreadsheet and presentation software
Reference	Encyclopedia or document reference material
Games	Software for pure entertainment
Video/Audio	Audio and video clips and webcam application

5. RESULTS AND ANALYSIS

Using the customised tool set, and based on the dimensions in Section 4.1: type of location, site, category of application, and characteristics of users, we explore relationships between categories. The following section presents and also discusses certain findings, both general (Section 5.1) and as they relate to research questions 1, 2 and 3, which are used as headers for Sections 5.2, 5.3, and 5.4, respectively. Research question 4 is addressed in the concluding section.

5.1 Locations, sites, applications and users

Table 2 lists the number of registered users per site, decomposed into male, female and unspecified, as well as the number of application hits by guest users and self-registered users, and the period over which data was collected. Table 4 lists the number of self-registered users per age group. The greatest numbers are in the 10-13 group (726 users) and the 14-17 group (749), with a high number aged 18-21 (583). This finding, which indicates extensive usage by secondary school learners, is most satisfactory.

Table 4: Age categories and number of registered users per category for the ten sites aggregated

	Description	Age range	Reg. Users
1	Typical junior primary school learners	6 – 9	57
2	Typical senior primary school learners	10 – 13	726
3	Typical secondary school learners	14 – 17	749
4	Post-school and tertiary level learners	18 – 21	583
5	Young adults	22 – 25	243
6	Older adults	26 – 60	279
7	The rest	0-5, 61+	389
	includes duplicate registrations		3026

Less pleasing data from Table 4 is that, from the subset of 10 sites, the logs record 3026 registered users of which 3026 – 2150 = 876 are duplicates. Of the 2150 unique names, only 1153 have a hit count of 1 or more, thus 997 user names were created and never used, suggesting either a lack of understanding of the purpose for creating a user name, or an inability to remember the user name and/or password combination.

With this background it was appropriate to drill down and determine how many registered users actually access their accounts. Figure 4 plots hit count (x-axis) versus number of users (y-axis). It indicates that a large proportion of users register a user name, but spend very little time logging in and using the account. By contrast, other users make extensive use of DD terminals. Notable examples follow:

- 500 users (almost 50%) launched only 1 application
- Over 200 of 1153 users (17%) registered 50 application launches or more (from a sample of users with a non-zero hit count), indicating extensive use.

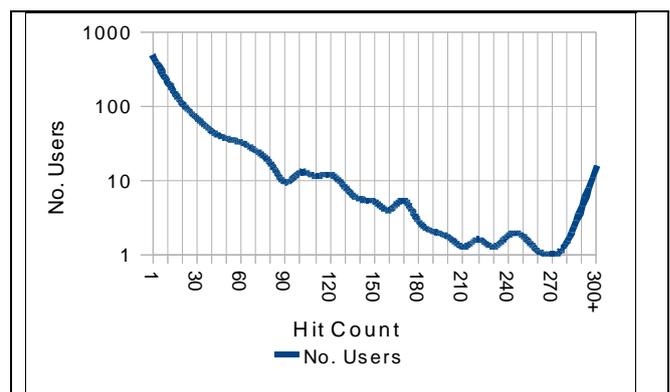


Figure 4: Number of users vs hit-count

5.2 What categories of applications are used by various age groups? (ResQ 1)

Figure 5 indicates the percentage of total hits per age group for each of the seven general application categories mentioned earlier, enabling comparison both of application usage within the same age group, and of differences in application usage between different age categories. We notice a downward trend

in the usage of games with increasing age, declining from over 40% of usage time for 6-9 year olds to 17% for users over 26. However, games remain the top- or second-most popular activity among all age groups. Education has low usage levels, less than 10%, except in the 26-60 group.

The low use of the office suite applications (OpenOffice word processor, spreadsheet etc) is notable. This is possibly due to the relative 'dryness' of these applications relative to games and video, but a further major issue is the absence of a printer at DD

terminals. Requests for printing services are frequently made by users, but the logistics of managing supplementary consumables in remote unsupervised areas have prevented the installation of printing facilities at any of the sites in this study.

Moreover, DDs are explicitly not intended to be clerical service points. Such usage would run counter to the envisaged fundamental purposes. The project may, however, require expansion to provide dual functionality of computer training and service delivery for sustainability reasons.

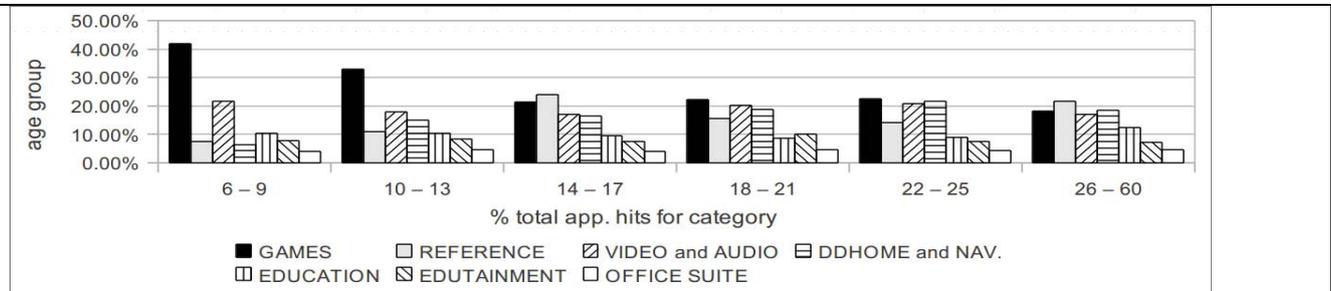


Figure 5: Categories vs age groups

Table 5: Per-site hit-count per age group

Location	Site Name	Age Groups														All age groups	
		6-9		10-13		14-17		18-21		22-25		26-60		Other		Total	%
		No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%		
1 School	Kwam_Hlonipa	0	0.0%	120	26.4%	171	37.6%	61	13.4%	73	16.0%	14	3.1%	16	3.5%	455	100%
	Vezebuhle	231	2.6%	3821	43.6%	1954	22.3%	901	10.3%	270	3.1%	275	3.1%	1303	14.9%	8755	100%
	Elandskraal	202	7.3%	910	33.0%	874	31.7%	117	4.2%	110	4.0%	411	14.9%	134	4.9%	2758	100%
3 Library	Emjindini	49	0.4%	1782	14.2%	6392	51.0%	1530	12.2%	406	3.2%	800	6.4%	1566	12.5%	12525	100%
	Kanya-mazane	0	0.0%	8	1.9%	74	17.6%	251	59.6%	70	16.6%	12	2.9%	6	1.4%	421	100%
4 FET College	LetabaFET2	2	0.2%	93	9.8%	28	2.9%	691	72.7%	82	8.6%	13	1.4%	42	4.4%	951	100%
	LetabaFET Giyani	0	0.0%	804	10.6%	427	5.6%	3205	42.2%	2385	31.4%	332	4.4%	445	5.9%	7598	100%
5 Cust. Care Centre	Msunduzi	0	0.0%	366	19.9%	369	20.0%	218	11.8%	424	23.0%	140	7.6%	324	17.6%	1841	100%
6 Fablab	SoshunguveFL	28	0.6%	1391	30.1%	1525	33.0%	1067	23.1%	95	2.1%	388	8.4%	134	2.9%	4628	100%
7 Informal Market	Kagung	0	0.0%	0	0.0%	402	21.2%	1141	60.1%	229	12.1%	127	6.7%	0	0.0%	1899	100%

Table 5 presents a useful overview of the age distribution at the various sites, as well as the extent to which the DD is being used (when related to the data collection periods in Table 2). For example, at Vezebuhle School, 16.5% of registered users are 18 or older, suggesting usage outside of school hours by older members of the community as well as by teachers. We also notice that the Fablab is frequented mainly by 10-21 age groups, Msunduzi Customer Care Centre has the highest number of registered users in the 22-25 group. Age groups with the highest usage of applications are highlighted in red in the table and reveal the following: the highest percentage usage at the first school is by the 14-17 age group, while at the second school the most extensive use is by 10-13 year olds. Interestingly, at the MPCC, highest usage is by 10-13 year olds. Library 1 shows the highest usage is by 14-17 year olds, while library 2, both of the FET colleges and the Informal Market show the highest usage is by 18-21 year olds. The customer care centre applications are mostly used by 22-25 year olds, closely

followed by users in the 10-13 and 14-17 age groups. The Fablab is predominantly used by the 10-13 and 14-17 year olds.

5.3 Does gender have an impact on extent of use and on application usage? (ResQ 2)

Figure 6 displays the percentage of males versus females accessing each of the general categories of applications (as a factor of all application launches). Table 2 (Section 4.4) indicates that, of total registered users, 1190 are male (77%) and 365 female (only 23%) and that only one female registered an account at the Kagung site. (However, at the popular Emjindini library site, the ratio of girls to boys was substantially higher than average, namely, 118:250). Hence it comes as no surprise to see that that the males in Figure 6 outnumber females by at least 3:1. While HITW sites in India also displayed usage behaviours of more boys than girls [5], the difference is even more marked in South Africa. The low usage by females is a

cause of concern. This study indicates that usage by females is considerably less than usage by males.

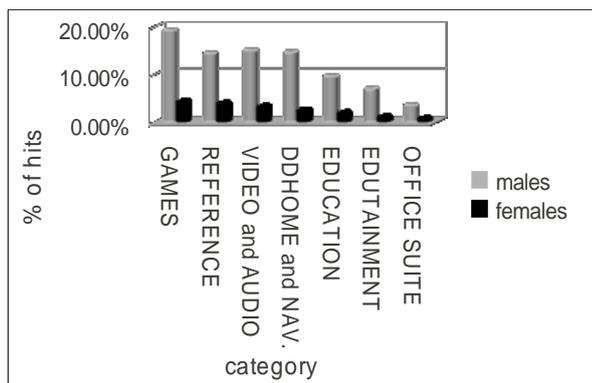


Figure 6: Percentage males and females accessing each category (as a percentage of total hits)

Not all users disclosed their gender, but calculations based on the column, 'Tot. Reg Users (m, f, u)' in Table 2, Section 4.4 indicate that the percentages of female users, calculated by $F/(F+M)$ in the ten sites respectively, are only: 26%, 19%, 25%, 32%, 19%, 21%, 21%, 31%, 16%, and 4%.

Figure 7 also plots percentages of males and females accessing each category, but as a percentage of just males, or just females respectively, attempting to determine if different genders show a preference for different general categories of applications. There is a statistically significant difference in the reference and DD Homepage categories. The difference in Edutainment bar is predominantly due to the popular 'TuxMaths' game, which is preferred more by males than females. Figures 6 and 7 are based on the values in the table below (Table 6).

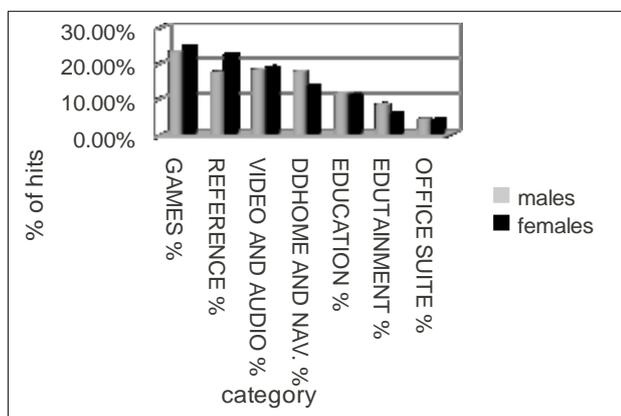


Figure 7: Percentage males and females accessing each category (as a % of total category hits per gender)

While useful trends and generalisation can be extracted from the categorised data above, a complete understanding of the usage patterns is only possible when referring to the composition (in terms of applications) of each category. Nevertheless, it is clear from Table 6 that for males, games are most popular followed by video and audio and the home page and navigation application. With females, games are seen to be the most popular, followed by reference material and video and audio content.

General Application Category	Usage of applications according to gender (M=male, F=female)					
	Totals per gender		% of app. usage per gender		% of total usage (M+F) of ALL applications	
	M	F	M	F	M	F
Games	6035	1396	23.5%	25.1%	19.3%	4.5%
Reference	4328	1247	16.8%	22.4%	13.8%	4.0%
Video and audio	4757	1060	18.5%	19.1%	15.2%	3.4%
DDhome and nav.	4650	768	18.1%	13.8%	14.9%	2.5%
Education	2527	525	9.8%	9.4%	8.1%	1.7%
Edutainment	2282	334	8.9%	6.0%	7.3%	1.1%
Office suite	1119	226	4.4%	4.1%	3.6%	0.7%
Total	25698	5556	100%	100%	82%	18%

Table 6: General application category usage

5.4 Which applications are underutilised by the intended target groups? (ResQ 3)

In Section 4.4 it was explained that, for the purposes of this research, the software applications available in DDs were categorised into seven general categories. To drill down deeper, an elaborated classification framework was proposed, involving 26 subcategories. Figure 9 and Table 7 are based on this more specific set of application groupings and allow more detailed usage analysis. Furthermore, Figure 9 clusters usage according to the following four types of location: libraries, secondary schools, FET colleges and general public places.

Some notable points from the hit-count study, calculated from the figures in Table 7, are:

- Games and pure entertainment: for users aged 6-9, 42% of the hits were for games and entertainment and for those aged 10-13, 33% were for games and entertainment. In older age groups there were notable downward trends in games usage; in each of the 14-17 and 18-21 and 22-25 groups the percentage was 22%; and for the 26+ age group, it dropped to 18%. Apart from Science education, education / edutainment applications were underutilised.
- In the age group 14-17, 16% of the hits were on Wikipedia. This may indicate use for school projects. Wikipedia was not used extensively by other groups.
- Use of file manager and home page: A user who knows exactly which application he/she wishes to interact with, accesses such an application via the menu. On the other hand, users wishing to explore available applications and functionality, must do so via the file manager and/or the home page. The DD home page also presents info on the background and purpose of the Digital Doorway. Both the DD homepage item and the file manager (navigation) item are in primary positions at the top of the 1st tier menu (Figure 8), and could thus be likely choices for new users exploring the content in an undirected way. However, hit counts on these pages show access to the file manager and home page increasing as a function of age as follows: For age groups 6-9, 10-13, 14-17, 18-21 and 22-25, respective hit-count levels are 6%, 15%, 16%, 18%, and 21%. This shows that focused searches and exploration reach their peak in the 22-25 age group. This is important, because the

22-25 year-olds represent users who are no longer at school and who are not based in an environment with a DD immediately available. They are making a voluntary decision to visit and use a DD. Furthermore, they are using the manager and home page to explore and select applications of their choice.

Table 7: Hit-count for specific categories (per age group)

Category	Age Category Hit-count						
	6-9	10-13	14-17	18-21	22-25	26-60	Other
video_audio, webcam	30	604	787	802	339	179	295
video_audio, science	3	106	114	61	38	14	34
video_audio, fun	74	839	969	750	367	180	332
video_audio, audiobooks	4	104	191	224	108	55	53
System, filemanager	11	794	964	825	379	202	263
System, ddhomepage	22	567	954	831	481	255	267
System, comms	0	29	83	57	23	8	4
Reference, wikipedia	7	470	1946	647	150	279	310
Reference, science	8	120	217	196	105	54	105
Reference, life skills/health	8	183	294	177	154	105	111
Reference, fun	3	34	71	55	29	8	13
Reference, books	10	134	253	221	86	61	79
Reference, agriculture	2	72	135	130	57	37	42
Office, openoffice	20	427	496	422	175	117	172
Games, orientation	112	1666	1106	666	275	168	529
Games, gamesother	101	1381	1500	1352	644	286	587
Edutainment localgames	0	99	178	349	9	70	15
Edutainment science	2	44	65	48	17	7	22
Edutainment maths	11	191	203	165	96	33	78
Edutainment language	5	34	45	65	27	15	22
Edutainment graphics	6	96	139	99	65	32	34
Edutainment gcompris	13	282	227	162	88	19	59
Edutainment computers	3	23	52	24	12	5	10
Education, science	40	757	849	565	286	200	365
Education, mindset	8	168	249	168	54	98	142
Education, maths	5	37	50	48	27	11	10
	508	9261	12137	9109	4091	2498	3953

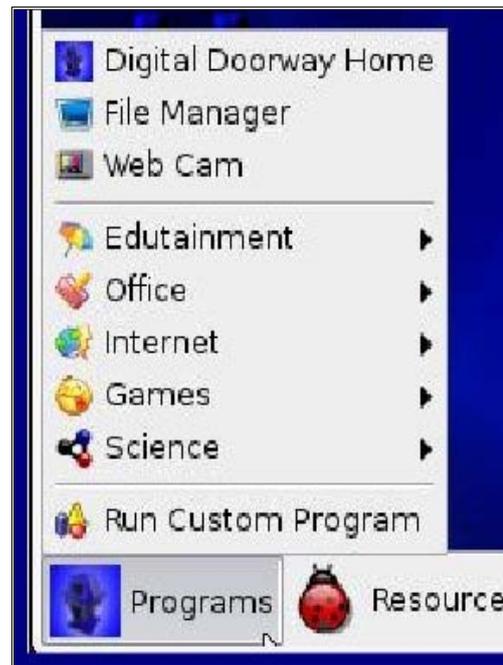


Figure 8: 1st tier menu

It is evident from the hits portrayed in Figure 9 that, despite the large selection of resources available on each DD, usage of just a few applications and application categories comprises the bulk of the hits. Games, which are launched from a 2nd tier menu, are the most popular selections. In addition, users at both schools and libraries tend to access and return to the webcam, fun video clips, the DD home page and Science education (predominantly interactive graphical simulations). Reference materials for agriculture, books, science, etc are little used. Edutainment of all varieties is used even less.

An interesting discrepancy between usage at libraries and at schools is the percentage of hits to the free encyclopedia, Wikipedia, which is accessed more than other reference media. Figure 9 shows high use of Wikipedia in libraries, which could be due both to a desire by library users for information acquisition and to the presence of a librarian to facilitate information retrieval. In future work, the issue of assistance by facilitators such as librarians and teachers and how this can work in conjunction with unassisted learning, will be investigated in on-site qualitative studies at selected sites. Usage in libraries is of great importance in analysing DD usage. Among the registered users (Table 2), the number of library users in the stereotypical sites selected for this study is second only to the secondary school users. Moreover, library visitors were almost twice as likely to access health information as school users.

Further figures (not included in this paper due to space restrictions) show specific categorised application usage by the various age groups. They portray a similar view to usage by location (Figure 9), indicating disappointingly low usage of edutainment, reference materials and office applications, all of which were deliberately placed in DDs for their value in education and training.

The values in these figures are actual hit counts and not percentages, and accurately depict the combined usage of Digital Doorways across all age groups and categories. High usage of Wikipedia by 14-17 year olds was noticed, this figure being heavily influenced by the fact that Emjindini Library has the largest base of registered users, over 50% of whom are in

this age group, and who use Wikipedia regularly. At other venues too (e.g. Gatang high school, not featured in this study), users are explicitly required by teachers to source information from the DD for school projects, a combination of unassisted and directed learning.

6. CONCLUSIONS AND FUTURE WORK

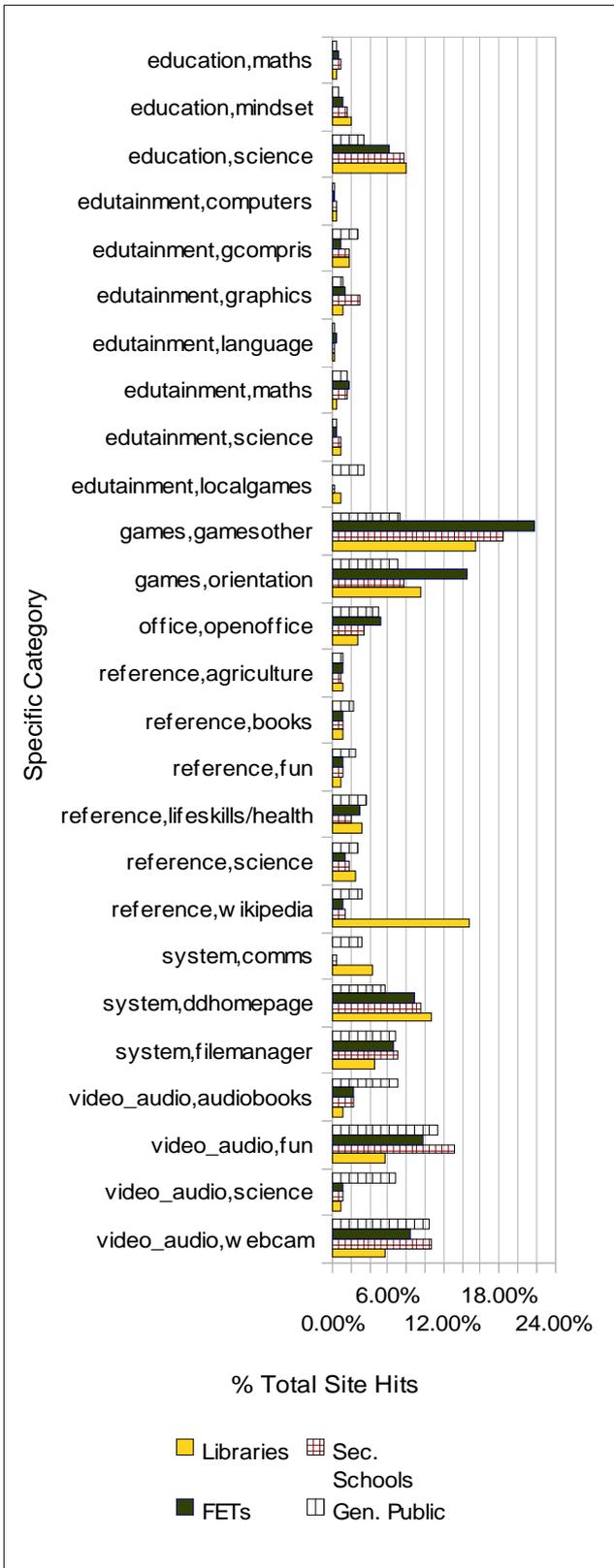


Figure 9: Specific application vs. location

This paper explained the South African Digital Doorway concept and discussed usage of its embedded applications with relation to aspects of user demographics and type of location. We analysed which categories of applications are being used by various age groups (Section), whether gender has an impact on application usage (Section) and which applications are being under-utilised by the intended target groups. (Section). This concluding section addresses the over-arching fourth research question, as well as providing recommendations, and suggesting topics for future research.

6.1 Lessons learned for future development, selection of applications, and presentation of applications (Res Q 4)

From analysis to date of the carefully-selected sample of log files, it emerged that DDs are being used by thousands of users from a variety of age groups, with the bulk of users being males between 10 and 21. Usage by females is approximately 25%, which is inadequate.

The DD experience is aimed to be self-directed and enjoyable for all community users. It results in valuable peer-assistance (noted in previous observation studies), implicit learning and computer literacy. However, it became clear that useful educational material and sources of knowledge enrichment are underutilised.

How should these issues be addressed? First, efforts could be made to explicitly encourage use of the direct educational material. Without changing the ethos of unassisted non-invasive learning, supportive signposting measures could be sought. Users might benefit from clear visual guides explaining the various types of content available, either in the form of an online document (immediately accessible on logging in) or eye-catching laminated posters attached to the unit itself, recommending use of particular applications and customised to the type of site. It is also important that the best-practice efforts used at particular venues such as Emjindini library (Section 5.4) should be explicitly documented and advocated to facilitators elsewhere such as school principals, teachers and community librarians.

Second, modifications could be made to existing under-utilised software, and new applications could be incorporated or developed, with a view to integration of educational aspects into game or audio-video presentation formats, which are the most popular. Reference material could be made prominent in the user interface to stimulate greater usage. Content choice and design would benefit from such efforts. In such processes, the findings of the Meraka group investigating DD usability, should be noted and applied. The position of items within the menu system influences the number of launches, therefore sequencing should be carefully designed to cater for first-time users.

Thirdly, qualitative research is required to complement the current work. Although this study has addressed only a few aspects of usage analysis, quantitative research has proved invaluable in revealing aggregated trends, patterns and occurrences, as well as usage behaviour across a representative sample of DDs. This data should, however, be used in conjunction with qualitative research methods to obtain the best possible picture of usage and to make suggestions for improving the users' experience.

6.2 Summary of recommendations

Following on the lessons learned, some specific recommendations arise.

- In response to the findings in Table 2, Section 4.4 and discussion in Section 5.3, that usage by females is low across the board, we recommend a campaign to stimulate interaction by females. (This would be in line with the SA government's efforts to promote science, mathematics and technology among girls and women.)
- Encourage greater interaction by the older generation (ages 26+) through awareness campaigns and more content targeted at them, paying particular attention to locations most frequented by adults (Table 5).
- With reference to the discussion in Sections 5.4 and 6.1, modify the content and develop new applications to stimulate user-interaction with educational resources.
- Without side-lining unassisted learning, facilitators should be trained in practical ways of enhancing utilisation of this valuable resource in their environments. Such training should emphasise the worth of educational resources and interaction by females (Sections 5.3 and 5.4). Moreover, educators and trainers should advise learners to access DDs in projects and in preparing deliverables.

6.3 Future Work

Following on the findings of the present study, possible future research directions are:

- Qualitative measures (as suggested in Section 6.1) such as interviews to find social reasons for the trends, patterns and outliers encountered in the quantitative analysis. Interviews should investigate: access by different age groups to the DDs; why users visit the DDs; how often they visit them; the levels of enjoyment and frustration experienced in usage; favourite applications and why they are favourites; and preference between single or collaborative usage.
- Studies to determine whether level of usage is translated into academic performance and workplace skills, for example, whether the Maths games improve school performance in the subject of Mathematics.
- Comparison of data from a variety of user interface styles, to ascertain to what extent usage behaviour is driven by directed behaviour and influenced by menu layout.
- Differentiation studies on how interactions at pure unassisted learning sites differ from behaviours at those sites where adult supervision plays a role. As mentioned in Section 5.4, this will be particularly investigated in library and school sites.
- A study at a location where printing facilities are hooked up to the system (see Section 5.2), investigating impact on usage behaviour, in particular office-related applications such as the word processor.

In all future ventures, however, the collaborative efforts of researchers and developers should be focused on making a positive and measurable difference to the lives of the poorest members of society. The provision of accessible infrastructure and relevant and engaging content combined with users' curiosity and desire for self-discovery should ultimately contribute to the acquisition of skills and knowledge vital to life in a digital world.

REFERENCES

- [1] Department of Education, 2009. Computer centres summary grid for ordinary schools. [Online]. <http://www.education.gov.za/emis/emisweb/statistics.htm> [Accessed July 2010]
- [2] Dangwal, R. 2005, Public computing, computer literacy and educational outcome: Children and computers in rural India, International Conference on Computers in Education.
- [3] Amiel, T. & Reeves, T.C. 2008, Design-based research and educational technology: Rethinking technology and the research agenda, Educational Technology and Society, vol. 11, no. 4, pp. 29-40.
- [4] Mitra, S. 2000, Minimally Invasive Education for mass computer literacy. [Online]. <http://www.hole-in-the-wall.com/docs/Paper01.pdf> [Accessed May 2010]
- [5] DeBoer, J. 2009. The relationship between environmental factors and usage behaviors at 'Hole-in-the-wall' computers. International Journal of Educational Development, Vol. 29, no. 1, pp. 91-98
- [6] Veeraraghavan, R., Singh, G., Toyama, K. & Menon, D. 2006. Kiosk Usage Measurement using a Software Logging Tool. Proceedings of the International Conference on Information & Communication Technologies for Development, at Berkeley.
- [7] Dangwal, R. & Kapur, P. 2009. Learning through teaching: Peer-mediated instruction in minimally invasive education. British Journal of Educational Technology, Vol. 40, no. 1, pp. 5-22.
- [8] HIWEL, 2007. Hole-in-the-Wall Education, Ltd. [Online]. <http://www.hole-in-the-wall.com> [Accessed May 2010]
- [9] Gush, K. 2008, Towards a more personalised user experience and better demographic data on the Digital Doorway public computer terminals. 5th Prato Community Informatics & Development Informatics Conference 2008: ICTs for Social Inclusion: What is the Reality? Conference CD. 27 October-30 October, Monash Centre, Prato Italy. December 2008.
- [10] Dangwal, R., Jha S. & Kapur, P. 2006, Impact of Minimally Invasive Education on Children - An Indian Perspective (colloquium), British Journal of Educational Technology Vol 37 No 2, pp. 295-298.
- [11] Mitra, S. 2003, "Minimally invasive education: a progress report on the "hole-in-the-wall" experiments", *British Journal of Educational Technology*, vol. 34, no. 3, pp. 367-371.
- [12] Gush, K., Smith, R., Cambridge, G. 2004, The Digital Doorway, Minimally Invasive Education in Africa. [Online]. http://www.digitaldoorway.org.za/publications/ICTinEducation_conference_paper- The Digital Doorway_March_2004.pdf [Accessed May 2010]
- [13] Smith, R., Cambridge, G. and Gush, K. 2005, Unassisted Learning – Promoting Computer Literacy in Previously Disadvantaged areas of South Africa. Proceedings of the WSIS Conference, 2005.
- [14] Gurstein, M. 2010, Let's talk about a digital transition. [Online]. <http://gurstein.wordpress.com/2010/04/02/let's-talk-about-a-digital-transition-rather-than-a-digital-divide/> [Accessed May 2010]
- [15] Oates, B.J. 2006, *Researching information systems and computing*, Sage.
- [16] Hilbert, D.M. & Redmiles, D.F. 2001. Large-scale collection of usage data to inform design. Proceedings of

the Human-Computer Interaction conference – INTERACT, pp. 569-576.

[17] Rahm, E. & Do, H.H. 2000. Data cleaning: Problems and current approaches. Bulletin of the Technical Committee on Data Engineering, pp. 3-13.

[18] [Olivier, M.S. 2004, Information Technology Research: A practical guide for Computer Science and Informatics. \(2nd ed.\) Pretoria: Van Schaik.](#)