USABILITY TESTING OF HAND HELD COMPUTING ON A CONSTRUCTION SITE

Sarah Bowden, Antony Thorpe and Andrew Baldwin
Arup, Department of Civil and Building Engineering, Loughborough University, Loughborough, UK, and Department of Building and Real Estate, Hong Kong Polytechnic University, Hong Kong, PR China
Sarah.Bowden@arup.com; A.Thorpe@lboro.ac.uk;A.N.Baldwin@lboro.ac.uk

SUMMARY

Unless current hand-held computers are found to be usable by site-based personnel the uptake of these new systems will be slow regardless the benefits available to these individuals and the project team as a whole. The technology to extend IT solutions to personnel in the field is available, but there is a preconception that site personnel are not IT literate and therefore will not be able, or willing, to take full advantage of the benefits that IT tools bring. This paper presents a methodology for assessing usability, describes the usability testing of hand held computers by site workers and concludes that this type of device will meet their needs.

INTRODUCTION

Computers are now sufficiently small to be carried casually and as such it is evermore common to find engineers with these devices. This provides the possibility of remote accesses to typically office bound analytical software by construction site staff, (Pilgrim, M. et al, 2002). In addition, the use of mobile computing can significantly improve the flow of relevant information among the project participants, (Magdic, A. et al, 2002). If hand-held devices are to become universally acceptable further research is needed on the acceptability or 'usability' of such devices. The primary aim of the usability evaluation was to determine how easy site-based personnel find hand-held computers to use. The second, subsidiary aim was to compare various devices that were already commercially available. Following a desk-based review of hand-held computers, four different types of hand-held devices underwent a series of usability tests. The tests were based on an accepted methodology that had been developed for general product user testing. Seventeen site-based personnel undertook four construction-based tasks on each device and then answered a series of questions about each task and the devices' physical attributes. All these staff were from the M6 Toll project, a major road construction project in the Midlands region of the UK. Over a two-day period these staff completed 68 tests. The results from these tests were then assessed together with the results of a questionnaire survey document, which was completed by each participant. The selection of the usability testing method is described together with the development of the testing procedure, its operation and the research findings.

HAND-HELD COMPUTERS FOR SITE USE

Mobile Computing hardware comes in many shapes and sizes. There are Personal Digital Assistants (PDAs), Pen Tablets, Handheld and even PDAs combined with mobile phones. Examples of these devices are shown in Figure 1. Previous studies of the use of mobile IT devices on construction sites have shown that users require the devices to satisfy the following criteria if they are to be acceptable for site conditions: the screen must be visible in bright sunlight and near darkness; the battery life should be at least 8 hours; the device must be able to survive being dropped from about 1m onto a hard surface; be able to be used in the rain; and be able to be carried in one hand, (Elzarka et al, 1997). The construction site is a tough environment with sunlight, rain, mud and heavy handling to contend with. But manufacturers are well aware of these constraints and are now providing hardware at various levels of 'ruggedness'. Rugged devices come at a typical cost premium of at least 50%. Rugged cases are also available for non-rugged devices.

Hand-held computers are now capable of running a range of software including: CAD applications; Collaboration Software; Data Capture; Project Management; and Discipline Specific Applications. Software suppliers are beginning to extend key collaboration features to mobile users in the field either through their mobile phones or other handheld devices. These applications allow site managers to view work programmes, and review daily, weekly or monthly tasks. CAD is now used on almost all construction projects to produce drawings for use in the field. However, although the drawings are produced electronically, they are generally printed out for use. This eliminates many of the advantages of electronic production, and reduces the opportunities for effective feedback from the field. Hand –held devices are capable of providing such a facility. Data Capture on site can be used to perform site safety audits snagging, quality inspections, resource Management, etc. Using a mobile device and the appropriate software almost *any* process that is currently performed using a clipboard and pen can be replaced.

SELECTION OF A USABILITY EVALUATION METHOD

Usability may be defined as: "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use."(ISO 9241-11: Guidance on Usability 1998). There are several different Usability Evaluation Methods, (UEMs), available for the evaluation of the usability of information and communications technology based products. These include: Cognitive Walkthroughs; Heuristic Evaluation; Usability Testing; and Pilot Testing. Different methods are suited to different design contexts and the time available to undertake the study. Pilot testing for example may take an extended time period and is best suited for prototype systems or extended evaluations before a consumer makes a significant capital purchase. (In the context of our research it was clearly inappropriate.) Table 1 provides a comparison of the other three UEMs considered. (This Table was adapted from that provided at www.userdesign.com) From this table it was concluded that Usability Testing was the most appropriate UEM to use for the purpose of this research.

Usability Testing was introduced in the late 1980s and rose to popularity in the 1990s (Wichansky, 2000). Gaffney (1999) defines usability testing as a technique for ensuring that the intended users of a system can carry out the intended tasks efficiently, effectively and satisfactorily. Rubin (1994) provides a complementary definition; "usability testing is the process that employs participants who are *representative* of the target population to evaluate the degree to which a product meets specific usability criteria". Usability testing encompasses both quantitative and qualitative analysis and as such tests can range from very large sample sizes to a single user. It aims to identify and rectify deficiencies existing in equipment prior to release. The intention is to ensure the creation of products that are easy to use, are satisfying to use, and provide utility and functionality that are highly valued by the target population. Due to the artificial situation that is created when conducting usability testing, successful tests do not provide 100% certainty that the product will be usable. However, usability testing when performed correctly can provide an almost infallible indicator of potential problems and the means to resolve them. It considerably reduces the risk of releasing an unstable or un-learnable product. Figure 2 is a model for conducting usability testing based on Rubin (1994). This was used as the basis for the testing of four hand-held computers.

DEVELOPING AND RUNNING USABILITY TEST

The aim of the usability evaluation was to compare various hand-held computers that were already commercially available and to find out how easy site-based personnel find these devices to use. The specific objectives of the usability tests were to: obtain a broad range of site-based personnel to act as participants; increase awareness of the types of portable I.T. devices that are available; identify the types of tasks that are best suited to hand-held computers; identify the functionality that site-based personnel would find useful; identify the views of site-based personnel about the use of hand-held computers in the construction industry; and determine which device the participants preferred and why.

Designing the test

In designing the test, feedback from a range of site staff was required. The target user profile was a range of staff including: Agents, Section Engineers, Site Engineers, Foremen and Inspectors. These personnel were to include persons of different ages, gender and experience who operated at different functions within the organisation. There was no restriction put on the level of IT experience of those involved in the tests but it was decided to test the devices with staff who had no previous experiences of using hand-held computers.

The experimental design focused on whether the users could undertake typical data collection and recording tasks using a range of hand held devices in a construction environment. In order to gain realistic results, the participants were tested whilst in their everyday working situation. Therefore the participants were required to conduct the usability tests whilst: standing up; working outside; and wearing site clothing i.e. helmet, coat, and boots and (if gloves were worn it was noted).

Careful consideration was given to the range of tasks to be included in the experiment. It was decided that the tasks should be representative of the information handling tasks that site-based personnel typically perform. They should also highlight different methods of data input and output; should use readily available software (financial and time constraints), should be able to be carried out on both the palmOS and Windows CE operating systems (device constraints) and should be intuitive and require minimal text input. A survey of the site staff was undertaken to contact potential participants and also to identify the information handling tasks they considered to be most important to evaluate. The results of this survey indicated that the document types which site-based personnel would find it useful to have access to/record in the field were drawings, data collection forms, correspondence, progress information and specifications.

The next consideration was the equipment to be used. The desktop survey showed that there are many hand-held computers available on the market, and new devices are appearing on a monthly basis. In order to test a variety of different devices, it was decided to obtain at least one device from each genre. However, due to the limited financial resources of this project these devices had to be obtained on loan resulting in the following devices being available for the usability tests: Rugged indoor screen hand-held PC (Itronix FEX21); PDA-Phone (Sagem WA3050); Rugged PDA with 28-key numeric keyboard (Symbol PDT8100); Rugged PDA with 16-key numeric keyboard (Casio IT700).

The devices obtained provided a range of different sizes, functionality, ruggedness and screen types. The Itronix device served to demonstrate the use of an indoor-specification screen outdoors. The suppliers provided the equipment on loan, with no restriction on the tasks chosen and with no constraints on the publication of the results.

Given the time and resources available to undertake the tests it was decided to limit the users to be tested to those from a single construction site. The opportunity arose to use staff from the M6 Toll project. This project was under construction by a consortium known as CAMBBA , (Carillion, Alfred McAlpine, Balfour Beatty and Amec). The users were recruited from the site staff following the questionnaire survey of potential users. This allowed the research team to collect initial data on the staff, their experience, their views and their willingness to undertake the hand held usability tests.

Setting up the test

The tests were conducted as part of a two-day event held at the headquarters of the M6 Toll project to increase construction staff's awareness of the potential of mobile communications. Each participant received User Instructions, a timeslot, and a participant questionnaire to reveal any additional details that could bias the results. The test procedure was as follows. Each device was set up prior to the tests with the required software and files. The order of testing each device by each participant was determined by random selection. Verbal and written instructions (displayed on a wall) were given, describing the usability tests, how to use the devices and the procedure for each task. Each participant performed the task in the defined order. After they completed that task they filled in the relevant question sheet. Once everyone had finished, the next task was described. No conferring was allowed, although questions could be directed to the test supervisor. Once all 5 tasks were completed, the participants received the next device. To ensure that the order in which the participants received the devices varied an appropriate swapping mechanism was used.

Running the test

On the days allocated for the testing of the hand held computers only 17 staff were available for the tests. Unfortunately, even with the provision of reserves, due to the participants' work demands three of the sessions only had 4 users. Instructions for the usability tests were read verbatim to each group in order that each participant was exposed to exactly the same conditions prior to the tests, the tester was not influenced by previous groups and adjusting the tasks accordingly, the instructions that were given were recorded for later use, and no points of instruction were omitted. The following tasks formed the usability test: a drawing task; a method statement task; a diary task; an inspection sheet task and a set of physical factors tasks. For the Drawing Task participants were asked to use PocketCAD to open a drawing and find out the width of a 'Family Room'. Then they had to imagine that they had actually measured this distance in the field (as-built) and that the dimension should be 6m. They were then asked to make a note of this on the drawing by using the drawing tools available to "cloud" the area, and write the correct dimension next to it. For the Method Statement task participants were asked to imagine that they were supervising the construction of the reinforced earth walls and were unsure how thick the backfill layers should be. They then had to access the method statement held on the device and from it find out how thick the backfill layers should be. For the Diary Task participants were asked to imagine that they were supervising a concrete pour and wanted to enter the details into their site diary using Microsoft Outlook. They were provided with the following activity details to enter: Date: 6th February 2002; Location:B360; Time: 10.00am – 16.00pm; Subject: Concrete Pour – East wing walls; Notes: Weather – fine. The concrete delivered 30mins late. The Inspection Test Sheet for catch pits for the M6 Toll project was converted into a form on each of the hand-held devices. Participants were asked to open and complete the form as if they were inspecting a catch pit in the field. This task evaluated the device itself regardless of the software on it. Participants were asked to consider the input methods, the screen and how comfortable the device was to use. After completing the tasks on each device the participants were asked to complete a set of questions that provided a comparison of their views on each device in terms of the physical factors, how easy the tasks were to perform on each, and which device they preferred overall. Participants were asked to consider the input methods, the screen and how comfortable the device was to use. After completing the tasks on all four devices the participants were asked to answer a set of questions that provided a comparison of their views on each device in terms of the physical factors, how easy the tasks were to perform on each, and which device they preferred overall. When the participants had completed all of the tests on all of the devices they were asked which device they preferred for each task and which device they preferred overall and why. This was followed by a videotaped group discussion to obtain further qualitative data.

Analysing the results

The usability tests were adjusted to reflect the actual number of participants and ensure so that the overall results were not compromised. With this relatively small sample (N<32) it is not statistically sound to generalise the results to the population (site-based personnel) however, the results collected can be used as guidance to the overall use of the devices.

On average the participants were able to complete 79% of the tasks with only a 10 minute training session and minimal instructions. The Rugged PDA (Symbol) was the preferred device for the Inspection Test Sheet (84% happy to use it); the Rugged PDA (Symbol) and the PDA-Phone (Sagem) were equally preferred for the Method Statement (79%) and the Diary (63%) tasks; and the PDA-Phone (Sagem) was the preferred device to use for the Drawing task (68%). Performing the tasks on any of the portable I.T. devices in order of preference were Method Statements (66%), Inspection Test Sheets (61%), Diary (53%) and Drawings (50%). Figure 3 shows device preference according to task.

The average results for each participant from the questions "Would you be happy to use this device for task x?" (Satisfaction score is 1 for yes and 0 for no) were sorted by age group, by job type, by I.T. experience and by hand-held computer experience. A one-way analysis of variance in each case showed that there was no significant variation across age (at 5% significance), job type (at 5% significance) I.T. experience (at 1% significance) or hand-held computer experience (at 5% significance) and therefore that differences in the means are simply due to sampling error. Previous experience of CAD and Microsoft Outlook does not have any significant influence over the satisfaction scores in those tasks (at 5% significance). Interestingly, contrary to commonly held beliefs, the Foremen and Works Managers were most enthusiastic in this sample with an average satisfaction score of 0.81 however, a one-way analysis of variance shows that at a 5% significance level the difference between the mean score for Foremen/Works Managers and the other job types is not significant.

In addition, participants were asked how useful they would find it to have access on site to the different types of information demonstrated by the tasks (scoring 1-5 with 5 representing 'very useful'). 'Method Statements and similar documents' were most useful (3.8), then 'Drawings' (3.6), 'Inspection Test

Sheets and similar documents' (3.5) and least useful was 'Diary' (2.9). Overall, the results were very encouraging with 15 (88%) of the participants confirming that they would be happy to use one of these devices for their work. The two (12%) that were not happy were Inspectors, in the 25-35 groups and the 45-55 groups one of who had not used a computer before and the other only for 1-2 years.

The methodology, developed for the usability testing of consumer products proved satisfactory for the testing of the equipment selected. The major problems with the method all related to the selection and availability of site staff to participate.

Factors that bias the results were also considered. Due to the time lapse between the initial survey and the trials, three participants (18%) had used hand-held computers. This was not considered significant overall, as these staff were not regular users of the devices. All of the participants were right-handed; this could be unrepresentative of the population). These tests are therefore unable to determine if these devices are suitable for both right and left handed people. To establish this a larger sample would have to be tested. Approximately equal numbers (three or four) of each job type participated. It was also noted that age groups were represented approximately equally.

There was a general level of enthusiasm for the future use of similar devices on a construction site. The users considered that, using such devices, information is less subject to the elements than other formats particularly paper. The information is easy to carry, rather than having a lot of paperwork 'filed' in the back seat of the pick-up truck. Collecting data electronically in the field and then synchronising it back to the site data network eliminate the tedious task of typing up notes when personnel return to the office.

The users considered that the devices could provide a useful reference tool so that personnel do not have to remember or predict what information they will need to view/record in the field. They could enable engineers to spend more time actually out on site. Data collected in the field will be more structured and consistent. There are also further benefits for the project team as a whole that result from having instant access to well structured data. Information collected in the field can be immediately passed on to other members of the project team. Data can be imported into other software packages. Data can be easily searched both for auditing purposes and for future knowledge management applications.

The participants identified the following barriers during the tests. Many found the stylus too small to handle with larger hands and potentially having to wear gloves too. It was considered that personnel might become too reliant on the device, such that if it were to break down they would have to go back to pen and paper and the necessary protocols would no longer be available. It was thought that the screen size available was not always practical for viewing drawings, and many would prefer to stick to A2 paper copies to carry out drawing-based tasks. Manual data input using either the stylus or the pop-up keyboard was found to be time consuming. (This indicates the need for manual input to be minimised through the use of drop-down menus and pre-written text.) It was thought that the costs involved in purchasing a device would outweigh the benefits gained. At approximately £1200 for a rugged device many participants thought that management would have to be convinced that purchasing these devices was worthwhile.

CONCLUSIONS

Usability Testing is a very useful method for evaluating the usability of information and communications technology based products. The model used in these tests proved satisfactory in all aspects. Through the involvement of representative users an understanding of usability of the product by its end users in the workplace is gained. Time, cost and accessibility of end-users were all factors that resulted in this testing having a sample that was too small to provide statistics that could be generalised to the target population. The results from this sample illustrate that there were no significant differences across job type in either preference for using the device or satisfaction with using the device. Also, previous hand-held computer experience and IT experience did not result in significant differences in the satisfaction scores for using the devices. The majority of the participants (88%) would be happy to use a hand-held computer on site, typical comments were *"Superb", "Very powerful", ". Definitely see an advantage".* However, the barriers of cost and training tempered these comments, and many participants reiterated the need for proof that the devices would be cost-effective, and that usable, useful applications would be available.

REFERENCES

Elzarka, H.M, Bell, L.C. and Floyd, R.L. (1997). Applications of Pen Based Computing in Bridge Inspection, Proceedings of the Fourth Congress in Computing in Civil Engineering. ASCE. June, p.327-334.

Gaffney, G. (1999) ‘Usability Testing’, *Usability Techniques Series*, www.infodesign.com.au.

ISO 9241-11: Guidance on Usability (1998) *Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability* International Organization for Standardization (ISO), 1, rue de Varembé, Case postale 56, CH-1211 Geneva 20, Switzerland.

Magdic, A., Rebolj, D., Cus-babic, N., & Radosavljevic, M. (2002), "Mobile Computing in Construction", in *International Council for Research and Innovation in Building and Construction, CIB w78 conference 2002*

Pilgrim, M., Bouchlaghem, N. M., Holmes, M., & Loveday, D. (2 A.D.), "Mobile Devices for Engineering Analysis", in *International Council for Research and Innovation in Building and Construction, CIB w78 conference 2002,*

Rubin, J. (1994) *Handbook of Usability Testing*, New York, USA, John Wiley and Sons, ISBN 0 471 59403 2.

Wichansky, A. (2000) ‘Usability testing in 2000 and beyond’, *Ergonomics*, VOL. 43, NO. 7, 998-1006.

Figure 1 Portable I.T. Devices

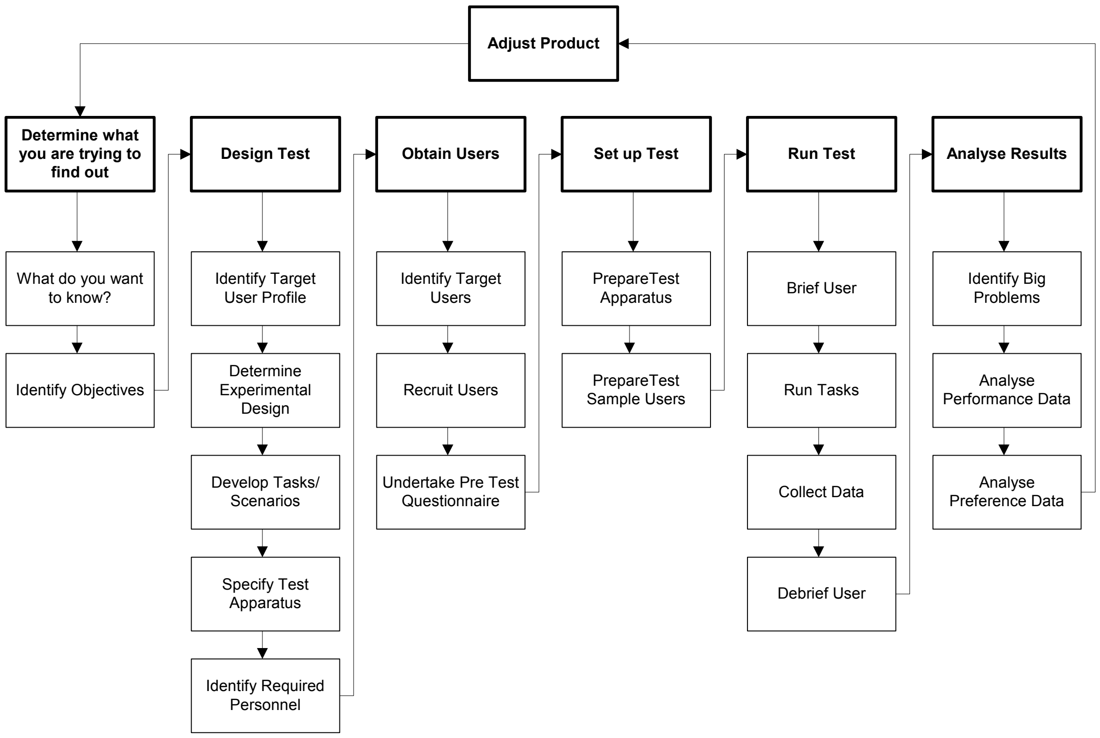

Figure 2 Usability Testing Procedure, (Model developed from Rubin, 1994)

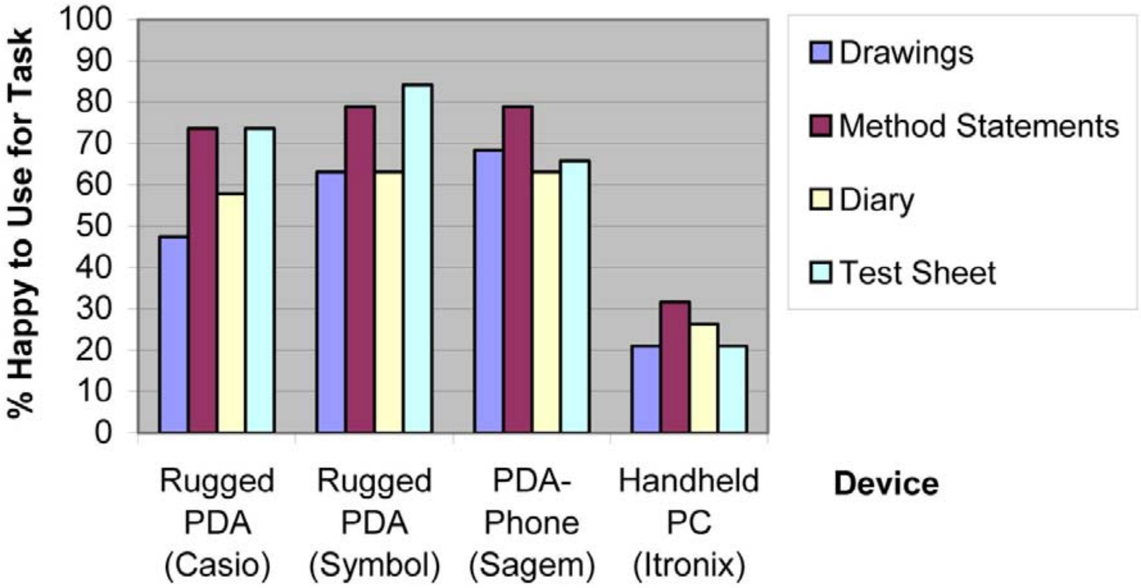

Figure 3 Device Preference According to Task

Name	Description	Synopsis	Advantages	Disadvantages
Usability Testing	Employs participants who are representative of the target population to evaluate the degree to which a product meets specific usability criteria by undertaking set tasks	• Uses representative users. • Uses scenarios and tasks.	• Uses representative users. • Can be conducted under real-world conditions. • Can discover “hidden” usability difficulties through un-prescribed user actions.	• Can be expensive and time consuming. • Minor usability difficulties can go unreported due to the semi-structured approach.
Heuristic Evaluation	HCI experts separately review an interface and categorise and justify problems based on a short set of heuristics (rules of thumb) / established usability principles.	• Uses short guidelines. • No scenarios or tasks. • Uses experts.	• Uses experts. • Gives multiple reviewers common rules to cite for justification of reviews. • Reasonably fast and cheap as it needs to be.	• The validity of Nielsen's guidelines (Mack and Nielsen, 1994) has been questioned and alternative guidelines exist.
Cognitive Walkthrough	A method, which fully utilises task scenarios to stress the user’s cognitive / problem solving process, checking to see if the simulated user’s goals and memory for action can be assumed to lead to the next correct action.	• Uses "information processing perspective" which puts the focus on the user's cognitive process and perception. • Uses scenarios and tasks.	• Puts the focus on the user. • May focus on known problem areas. • Recognition of user goals. • Uses the software developer.	• May be tedious. • Tries to make the designer the user (requires considerable commitment). • Inherent bias because of task selection. • Only addresses cognitive/ease of learning issues.

Table 1 Comparison of Usability Evaluation Methods