INTERNATIONAL CONFERENCE ON ENGINEERING AND PRODUCT DESIGN EDUCATION
8 & 9 SEPTEMBER 2011, CITY UNIVERSITY, LONDON, UK

A USABLE EVALUATION TOOL FOR DESIGNERS

Andree WOODCOCK[1], Simon FIELDEN[2] and Richard BARTLETT
[1]Coventry School of Art and Design, Coventry University
[2]Health Design Technology Institute, Coventry University

ABSTRACT
Evaluation is an essential, yet largely overlooked component in design education. Although a more user centred, inclusive approach to design is now advocated, practicing designers may not have been trained in the most appropriate ways to evaluate their designs. Reasons for this may include lack of resources and time available in the curriculum, lack of experience of lecturers in evaluation methods and a curriculum which emphasizes design production at the expense of evaluation. Without such evaluation, iterative design may only be informed by internal critical peer review. With a wider understanding of diversity and the need to design for an increasingly wide range of users there is a greater need to understand user requirements and evaluate products with representative end users.
A survey of SME developers of assistive technology products showed that they needed support in the selection of the most appropriate evaluation methods; that they may not have had much previous experience of evaluation, relied on a limited set of evaluation methods and were dependent on third parties gathering information for them. Based on previous experience of the development of paper and computer based design support tools and the teaching of research methods courses to designers, a decision support system was developed to guide the designers of assistive technology products in the selection of the most appropriate evaluation methods.

Keywords: Assistive technology, usability testing, evaluation support, ergonomics

1 INTRODUCTION

Young design students find it difficult to design for those unlike themselves. Usually they have had limited experience or opportunities to engage with people outside their peer or familial groups. This means that in the early stages of their design education they may design for themselves, or people like themselves as an approximation.
Previous authors have noted the growing requirement for student designers to be provided with methods and tools to enable them to gather information for themselves about end user populations. To this end, design textbooks, such as [10] include sections on task analysis and other ergonomics methods. [9] has produced a book on human factors (ergonomics) specifically aimed at design students and designers. More recently 'designer friendly' tools and methods have been developed (e.g. by IDEO) to encourage greater understanding and participation of end users in the design process. Such methods enable the designer to understand user requirements in a more relaxed, integrative manner. Frequent criticisms leveled at more traditional ergonomics tools and methods were that they required the designer to step outside of the design process, restricted creativity, were time consuming and of limited added value [11]
Vigorous efforts have been made (e.g. in Coventry University's undergraduate industrial design courses, [12]) to integrate ergonomics into the curriculum, rather than just presenting it as separate lectures, the content of which may be difficult to understand for those with non science backgrounds and of little immediate use. Undergraduate designers will work through scenarios, personas, storyboarding, use simulation suits, undertake product reviews and observe and interview members of their target audience in order to gain insights into needs and requirements. Much attention is placed on the front stage of the design process. As the design proceeds attention is focused on interaction, interfaces and the user experience, and user anthropometry – ensuring a goodness of fit between the user and the end product.
Evaluation of the final product with representative end users may not be given the same degree of attention. For example final concept design ideas may just be evaluated in design crits, or with clients,

as opposed to representative end users. Ethical review processes may impede the level of involvement of real end users, either through observation or interviews. This is especially the case for designs of assistive technology (AT) products requiring input from patients or medical practitioners, where a prolonged ethical review process, adapted from clinical practice, may restrict students' engagement with representative end users. For those going into design practices specializing in the new markets opening up for AT, this may mean that there is a shortfall in their levels of experience,

At another level, design is a practice based discipline. The emphasis is on the creation of form and studio based work. Student designers may attend lectures on evaluation and research methods as part of their studies, but unless the methods are designer friendly and/or directly applicable to their current project they may not be used or practiced. The generation of concept vehicles and products necessarily means that much evaluation is focused on storyboards and evaluation of physical mock ups.

Evaluation at all stages of product development, with representative end users is important in the creation of designs that please, delight and satisfy the customer, that have a high degree of safety and enable tasks or goals to be achieved effectively and efficiently. However, students exiting university may have had little exposure to designing and conducting evaluation, and little knowledge of how to select the most appropriate method for a particular stage of the evaluation and user type. It may be argued that such evaluation is not the responsibility of designers. However, SMEs (Small and Medium Enterprises) may have to rely on in-house staff to conduct such evaluations, to quickly assess whether a product is going to meet the requirements of its end users. Such evaluation is vital if AT products are going to benefit end users, and be accessible by their target user groups.

2 USE OF EVALUATION METHODS BY ASSISTIVE TECHNOLOGY SMES

To determine how evaluation is currently conducted in SMEs involved in the design of AT products, eight semi structured interviews were conducted to discover whether additional support is needed in for designers. The companies included a hospital based medical technology company, a company designing circuit boards for medical applications, a design consultancy, young company specializing in cognitive technology, design manufacturing company, a world leader in the design of telecare and assistive technology and a gait analysis SME. Product end users included commissioning organizations, monitoring managers, service telecare providers, housing associations, local councils, operators of telecare services, carers and care recipients.

All companies asserted that they were committed to user engagement throughout the design process from pre brief through to after sales follow ups, and cited instances of this. However, there was a reliance on informal qualitative methods (such as informal interviews, user forums and focus groups) and steerage from after sales support teams. Worryingly in some cases evaluation was just with representatives of end users (rather than actual end users), with no formal recording made of an evaluation. Problems associated with the involvement of users included:

1. Finding and accessing representative end users with certain medical conditions.
2. Managing end users expectations, especially in the early stages of concept development *'Users have to understand that the prototypes may not work*,'
3. Relationship management and education - building trust with end users, so that the right level of information is provided, design limitations understood and users trust designers not to 'steal ideas'
4. Costs of running evaluation and management of user groups
5. Perception of the value of end user contributions '*I can't think of an innovation that came direct from an end user'*, or '*Well that's pretty difficult; we are dealing with elderly and disabled section of the population so some of them are extremely bright, but the vast majority are being looked after for one reason or another and the very fact that they' re being looked after implies that they wouldn't necessarily contribute a great deal to the design process. When you're dealing with people who are already not in the best state of health then that type of feedback is not really forthcoming*.'

The interviews showed that SMEs appreciated the needs for continual user involvement, engaged with different categories of end users, evaluated a wide range of factors that aesthetics and wanted to build relationships and dialogues with end users and clients. However they relied on a few, qualitative evaluation methods and did not necessarily collect, record or analyze the data, and found it sometimes hard to reach end users.

There may be a gulf between methods used by SMEs and those developed by researchers. The SMEs believe that their current methods were sufficient, so introducing new ones may meet with resistance.

However, alternative methods may provide richer material and opportunities for greater engagement. Therefore, there is a need to both inform SMEs about other methods and to provide guidelines on how to design user experiences that are valid, robust and reliable.

3 DEVELOPMENT OF THE USER TESTING TOOLKIT (UTT)

UTT is a support system to assist designers (especially of Assistive Technology products) in the selection of the most appropriate evaluation method for their given circumstance, taking into factors such as user characteristics, stage of the design process, resource availability, context of use). As a learning tool, UTT takes designers step by step through the questions that need to be asked when considering a usability evaluation; it provides a short description of each method in the database, and links to where more information, provides alerts about methods where necessary (for example if resource intensive), or when accredited evaluation service providers should be brought in, and overall advice on how to conduct an evaluation. The methods included in the database are shown in Figure 1. and screen shots of the system in operation are shown in Figures 2-3.

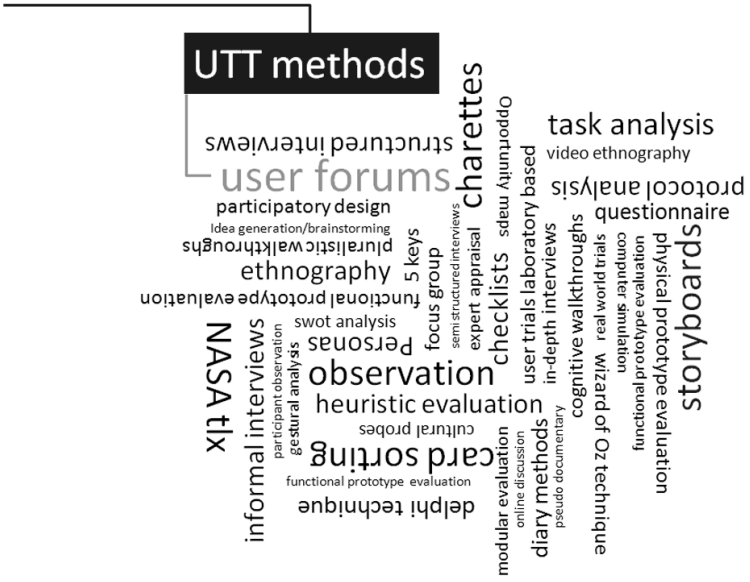

Figure 1. Evaluation methods included in the database

Most of this information may be included in undergraduate courses. However it may not feature prominently or be in an accessible form. Additionally designers may find the topic of little interest or difficult to understand, especially when quantitative methods and statistical tests are overemphasized. The need for systems such as UTT, have been recognized in the field of software design. [1] (p3253) commented, that difficulties associated with software evaluation required ' *A computer aided method selection system, which compares the general conditions of and demands on the evaluation (e.g. finance budget, target criteria, user participation and many more) with the characteristic attributes of the methods and suggests an optimised selection of evaluation methods (mixed method) would seem to be advantageous*.'

Previous research [9] has shown that such systems need to be fast, easy to use, provide the right level of information to enable action to be taken, designer/user friendly (avoiding jargon and technical terms), provide added value, not require duplication of work or extensive form filling, not make any preconceptions about the design process and be flexible enough to accommodate a wide variety of products and solutions. Added to this is the need to increase the system user's awareness of the sort of issues to consider when designing evaluation studies. It is hypothesized that this implicit aim will make the UTT attractive to design students and their lecturers, especially since the system can be used to support the design of any product and at any stage of the design process. The system runs as an independent executable, with a simple question and answer interface, which takes about 15 minutes to complete. Its output provides the top 9 most appropriate research methods based on the answers provided.

Factors which need to be considered in the design of evaluation studies were collated from the literature [e.g. 2, 3 and 4] and the experiences of the project team. This resulted in a set of general questions relating to stages of the development process, purpose of the evaluation, resource availability, end user accessibility, form of the product to be evaluated. Additional questions focused on disability, tasks being supported (e.g. communication, movement), end user group. Sets of potential answers were developed for each question, and their relevance checked by the project team. For example: the question relating to the length of time available to conduct the evaluation might have the answer set: one day, week, month, three months, longer. A further review produced a set of over 40 candidate research methods which could potentially be used to evaluate assistive technology products. Experienced researchers mapped the suitability of each research method on to each response. The importance of each group of questions (e.g. user characteristics, stage of the design process) was weighted along with the relative importance of the question in that category. In some cases a particular answer may mean the exclusion of a research method altogether. For example, if the product being developed relates to personal hygiene it may not be appropriate to use an observational study.
The UTT is a standalone application, programmed using C-sharp in Microsoft Visual Studio. Its development occurred through the close working of the programmer with a human factor's expert. An iterative process allowed rapid changes and developments, e.g. enabling the testing of different weighting algorithms. Major changes in the interface and functionality were decided upon as a result of running the program and reflecting on the system performance. Provision has been made to add additional questions, response categories and methods and relative weightings can be changed

4 EVALUATION OF THE USER TESTING TOOLKIT (UTT)

The questions and responses were assessed twice by the project team. In attempting to be comprehensive and precise, the language used in some of the questions alienated the end user designers e.g. when using Maslow's hierarchy of needs to categorize the product type. As the system aims to be applicable to a wide variety of products/services/systems and devices it is difficult to find the right terminology to cover all instances (for example the term product, solution, system have all had their advocates, and none satisfy everyone). In some cases questions seem repetitive, even though they tease out slightly different factors. This can be confusing and frustrating for designers. A second review led to the rephrasing of many questions.
Version 1 of the UTT was evaluated to confirm the appropriateness of the methods when compared with those actually used in evaluation studies. A good match between the suggestions derived from UTT and those used in the evaluations led to some confidence in the weighting scales.
A usability study was conducted with 7 representative SMEs to identify usability problems and suggest improvements. The SMEs were engaged in developing discrete AT products - a walking stick, independent living aid, exercise equipment for disabled users, heath monitoring devices, mobility aids for children, environmental controls.

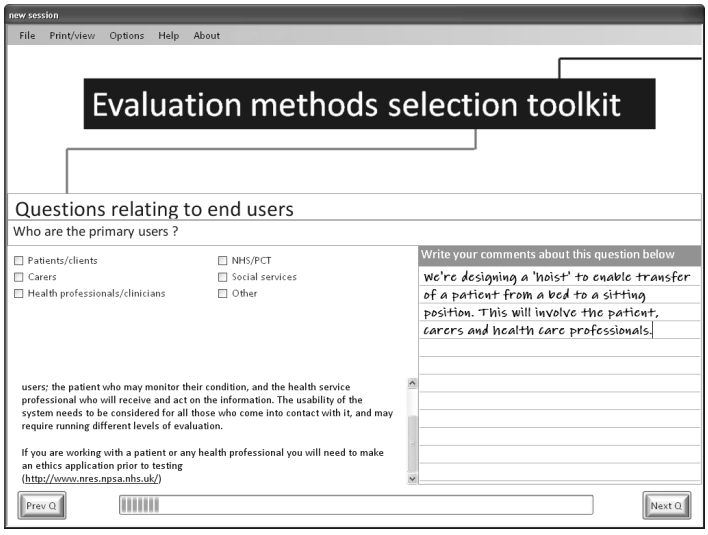

Figure 2. Example of the Q/A interface which forms the heart of the Toolkit

Participants described the product they were thinking of evaluating and then worked through the system. Walkthroughs [5] provide an indication of how easily users can perform tasks after little or no training. Observational studies can identify mismatches between the way in which designers think a system should work and users` actual experiences. Figure 2 shows the Q/A interface developed as a results of user feedback.

Participants were asked to comment on the usability of the toolset as they used it. Such verbal commentaries [6] reflect what participants are thinking and can help in understanding user interactions [7]. Instances were identified where participants experienced difficulties operating the toolset. Breakdown analysis [8] was used to rapidly identify and classify problems such as mismatches in how tasks should be carried out, misunderstandings of the terms used, layout of information and type of feedback.

All participants understood the system, and in most cases felt that it had generated appropriate research methods. However the terminology on many of the questions was still judged to be unfamiliar or inappropriate. Suggestions were given and where possible wording changed, additional user help and explanations were provided. Suggestions were made for optimizing the position of buttons, including more visual feedback to prompt user actions and provide a progression toolbar. Although no comments were provided about the need for system help, SMEs would need help defining their user groups, setting up an evaluation study and understanding different research methods. Figure 3 shows the results of the analysis, including the star system based on the scores, alerts where necessary, and a short description of the methods. Different levels of information can also be printed out.

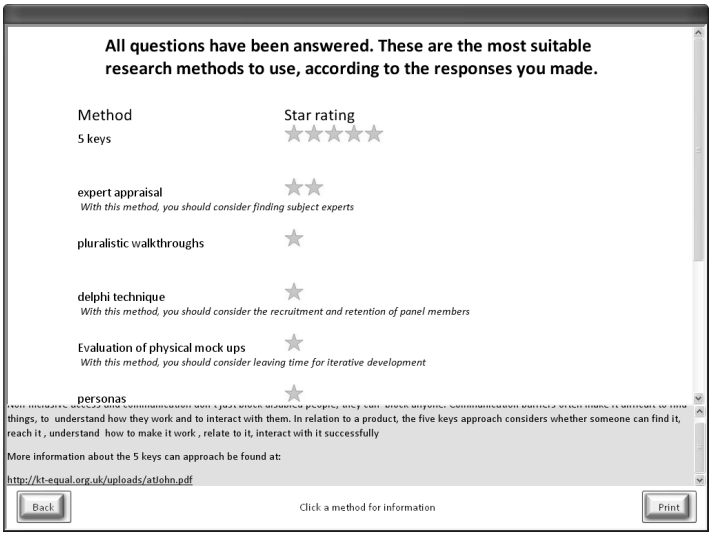

Figure 3. Results screen, showing the most appropriate methods

The prototype system kept the cumulative scores on the screen at all times, so the effects of answering a question could be seen on the order of the research methods. It was felt that this might be of interest to SMEs and promote learning. It was not. Participants were only interested in the final scores and felt that the provision of superfluous information cluttered the interface and biased their answers. The SMEs requested additional functionality in which more detailed information was provided for the evaluation, such as the questions which should be asked. As usability sessions have to be tailored specifically to the product and the users in question, it is not always possible or wise to provide generic questions.

5 FURTHER WORK

The current release of the toolkit is available on request to the authors or from Coventry University. Future developments will include release to students to test the educational benefits and usefulness of the system, and to use the system as a means of providing training to designers.

ACKNOWLEDGEMENTS

The research was supported by grants from the Technology Strategy Board and the Engineering and Physical Science Research Council

REFERENCES

[1] Karwowski, W. (Ed.) *International Encyclopedia of Ergonomics and Human Factors,* 2006 (2nd Edition, Taylor & Francis)

[2] Lumsden, J. (Ed.) *Handbook of Research on User Interface Design and Evaluation for Mobile Technology.* 2008 (Idea Group, Inc)

[3] Baber,N. and Stanton,C. Factors affecting the selection of methods and techniques prior to conducting a usability evaluation, in P.W. Jordan (ed.), *Usability Evaluation in Industry,* 1996 (Taylor and Francis, London)

[4] Roto, V., Obrist, M., Kaisa Väänänen-Vainio-Mattila,K., User Experience Evaluation Methods in Academic and Industrial Contexts, downloadable from http://wiki.research.nokia.com/images/9/9a/RotoObristVVM-CHISIGmethods.pdf. accessed 17th Sept 2009

[5] Polson, P. G., Lewis, C., Rieman, J., and Wharton, C. Cognitive Walkthroughs: A Method for Theory-Based Evaluation of User Interfaces. *International Journal of Man-Machine Studies* 36, 1992, 741-773

[6] Ericsson, K. A., and Simon, H. A. *Protocol Analysis: Verbal Reports as Data.* 1993 (Cambridge MA: MIT Press)

[7] Byrne, M. D., John, B. E., Wehrle, N. S., and Crow, D. C. *The Tangled Web we Weave: A Taskonomy of www Use.* CHI`99, 1999 (ACM Press)

[8] Woodcock, A., and Scrivener, S. A. R. `Breakdown Analysis`. In *Contemporary Ergonomics.* ed. by McCabe P. 2003, (Taylor and Francis; UK)

[9] Garner S. *Human Factors*, 1991 (Oxford University Press)

[10] Baxter, M. *Practical methods for the systematic development of new products*. 1995 (Chapman and Hall, London)

[11] Woodcock, A. and Galer Flyte, M. Supporting the integration of ergonomics in an engineering design environment, *Tools and Methods for Concurrent Engineering*,1998, 21- 23rd April, Manchester, England, 152-16

[12] Mackie, E. Ergonomics Teaching within Industrial Design; an Evaluation of Evidence of Understanding .*Proceedings of Contemporary Ergonomics 2011*, 'Contemporary Ergonomics 2011'. Held 12-14 Apr 2011 in Lincolnshire, UK