Nov 17th, 12:00 AM

Beyond Desktop Usability: Web Site Usability and its Evaluation.

Stephen AR Scrivener
Central Saint Martins College of Art and Design

Chao-Yang Yang
VIDe Research Centre

Andrée Woodcock
VIDe Research Centre

Beyond Desktop Usability: Web Site Usability and its Evaluation.

Stephen AR Scrivener

Central Saint Martins College

of Art and Design

Chao-Yang Yang

Andrée Woodcock

VIDe Research Centre

Understanding of Human-Computer Interaction (HCI) has progressed to the point where there seems to be broad consensus about what constitutes usability and how it might be evaluated (Brinck et al., 2002). In the last fifteen years there has been rapid development in the World Wide Web (WWW), which has grown to be an enormous information resource that users surf and mine for work, leisure, entertainment and transaction purposes. In the early days of the WWW, many web sites and pages were poorly designed. The rigour of conventional human-computer systems design (i.e., of desktop word processing, text editing, database applications) appeared to be lacking and it was not long before researchers sought to transfer HCI usability knowledge into the WWW domain.

The assumption that web site usability and HCI usability are the same has underpinned this transfer (Nielsen, 2000 and Brinck et a.l, 2002 etc), as reflected in the many web site evaluation guidelines and instruments that characterise web site usability in terms of learnability, efficiency, memorability, errors and satisfaction (Rubin, 1994). However, since HCI understanding largely concerns personal desktop computer systems designed to support users in specific work related individual tasks, it is quite likely to be incomplete or inappropriate with regard to web site usability. For example, conventional HCI systems are characterised by task directed activity. However, surfing the web is not essentially a task directed activity and users may access a web site purely to obtain intellectual and emotion gratification (Spool, 1999). Furthermore, typical desktop applications are used in a non-competitive context. For example, a desktop system usually has only one word processing application, whereas there is typically a set of alternative web sites that the user can choose from. Therefore, it is quite possible that what determines user satisfaction is different in each context. Furthermore, it is also quite possible that web site designers conceptualise usability differently from web site users. This is because designers are likely to be influenced by desktop focussed HCI knowledge and design values, whereas the users' view of usability will be determined largely by their experience of using web sites.

With this in mind, an Internet questionnaire survey was conducted in order to explore whether web users' and web designers' conceptions of usability differ, and whether these conceptions are encompassed within conventional HCI usability.

The primary results of the survey were that:

1. The users and designers had different viewpoints. For example, the users wanted web sites that provide diverse functions and information, whereas the designers thought web sites should be designed thematically. Similarly, the users valued information content over visual design, the designers vice versa.
2. Usability goals, such as stickiness (i.e., a web site's ability to engender user attachment to it), attractiveness and likeability were found to be as important to web site users as conventional HCI usability aspects, such as learnability, etc.
3. These attributes were found to be associated to specific web site features, such as navigation.
4. Users' were found to have very limited appreciation of the full capabilities of their favourite and most used web sites.

These results indicate that web site usability differs in important respects from conventional HCI usability and suggest that new evaluation methods are needed to accommodate the needs of web site users. This paper describes the internet questionnaire, the results obtained and the design and testing of a new web site evaluation method designed to both encompass an expanded conception of web site usability and to bridge the gap between designers' and users' conceptions of web site usability.

BEYOND DESKTOP USABILITY: WEB SITE USABILITY AND ITS EVALUATION

C-Y Yang[1], S A R Scrivener[2] and A Woodcock[1]

[1]The Design Institute, Coventry School of Art and Design, Coventry, UK
[2]Central St Martins College of Art and Design, London, UK

Contact email: s.scrivener@coventry.ac.uk

Introduction

Understanding of human-computer interaction has progressed to the point where there seems to be broad consensus about what constitutes usability and how it might be evaluated (Brinck, Gergle and Wood, 2002). In the last fifteen years there has been rapid development in the World Wide Web, which has grown to be an enormous information resource that users surf and mine for work, leisure, entertainment and transaction purposes. In the early days of the WWW, many web sites and pages were poorly designed. The rigour of conventional human-computer systems design (i.e., of desktop word processing, text editing, database applications) appeared to be lacking and it was not long before researchers sought to transfer HCI usability knowledge into the WWW domain.

Underpinning this transfer is the assumption that web site and HCI usability are the same (Nielsen, 2000; Brinck et al, 2002 etc), as reflected in the many web site evaluation guidelines and instruments that characterise web site usability in terms of learnability, efficiency, memorability, errors and satisfaction, i.e., LEMES (e.g., Rubin, 1994). However, since HCI understanding largely concerns personal desktop computer systems designed to support users in specific work related individual tasks, it is quite likely to be incomplete or inappropriate with regard to web site usability. For example, conventional HCI systems are characterised by task directed activity. However, surfing the web is not essentially a task directed activity and users may access a web site purely to obtain intellectual and emotion gratification (Spool, 1999). Furthermore, typical desktop applications are used in a non-competitive context. For example, a desktop system usually has only one word processing application, whereas there is typically a set of alternative web sites that the user can choose from. Therefore, it is quite possible that what determines user satisfaction is different in each context. Furthermore, it is also quite possible that web site designers conceptualise usability differently from web site users. This is because designers' are likely to be influenced by desktop-focussed HCI knowledge and design values, whereas the users' view of usability will be determined largely by their experience of using web sites.

This paper describes the results of a research programme designed to:

1. Identify the limitations of existent methods for web site usability evaluation.
2. Investigate the difficulties experienced by web designers in the application of usability data to the design or redesign process

3. To develop a method that overcomes the limitations of the methods considered in 1, while meeting the needs of designers better
4. Assess the usefulness of the proposed method for web designers

The following steps were taken to achieve objectives:

1. The limitations of current website evaluation methods were explored by using a representative sample of current usability evaluation methods to evaluate an existing web site. The analysis of the usefulness of these current methods revealed their comparative strengths and weaknesses, informed by the data obtained from an interview with the web designer that explored the intelligibility and usefulness for design purposes of the information obtained by each method.
2. An on-line questionnaire survey was conducted to explore potential differences in designers/ users perception of web usability, the actions web designers took upon site launch, and how they dealt with user feedback. This information was used to inform the development of a new framework for web site evaluation.
3. A new method was designed to overcome the problems identified from Stage 2, such as differences in value systems between users and web designers and designers' difficulty in understanding problem reports.
4. The method was applied to a new site to test its comprehensiveness and web site designers' attitudes to the design information generated using it.

In the following sections, we outline each of these stages and the results obtained, giving particular attention to Stage 4, i.e., evaluating the method.

The applicability of HCI methods to web site evaluation

As Brinck et al. (2002, p14) observe that, "A usability method is any technique you use to create a design from a user-centred perspective". Nielsen (1993) conducted a comparative study of the primary types of usability method and reported the advantages and disadvantages of each method, recommending that evaluators should not rely on a single usability method to the exclusion of others. Notwithstanding the fact that the best results can be achieved by combining the strengths of different classes of method, such methods have been found to be limited. For example, they can fail to provide useful data for design because they do not correctly define the problems that the users experience when using a web site (Berkun, 2001). Additionally, the data from users is sometimes not reliable (Spool, 1999), e.g., some users might equate appearance with usefulness and usability, or users may not be able to predict or adequately explain/classify the problems they have. Nevertheless, current HCI evaluation methods have been applied web usability evaluation, although these were not originally designed for this purpose. However, web sites are different to conventional personal computer applications and hence current HCI web site evaluation methods may be limited in additional ways to those currently reported when applied in the Internet context. To investigate this possibility we undertook a web site evaluation employing the following methods.

Observation - In order to calculate the frequency of correct task completions, task completion time, use of commands, and frequency of user errors, we

conducted video (Preece, 1994, p620) and protocol analysis (Gould and Lewis, 1985; Nielsen 1992, 1993). Protocol analysis adds an extra dimension to video and other forms of information gathered by addressing the cognitive activity underlying the user's physical behavior, e.g., the execution of a particular task; the user's reaction to breakdowns and system error messages, and so on (Preece, 1994).

Breakdown Analysis (Urquijo, Scrivener, and Palmen, 1993) is a method for evaluating "breakdowns" in user-computer interaction.This method provides data for systematically identifying what, when and why problems occur.

Meaning in Mediated Action, i.e., MIMA (Bourges-Waldegg, 1998; Bourges-Waldegg and Scrivener, 1998, 2000) is designed to assist designers in understanding how representations mediate the actions of culturally diverse users and how to tackle culturally determined usability problems. It is very important that a web site is understandable to users from different cultures.

Website Analysis and Measurement Inventory, i.e., WAMI (Kirakowski and Cierlik, 1998) evaluates user' satisfaction with a web site in terms of its attractiveness, controllability, efficiency, helpfulness, and learnability.

The UNITE web site, which provides information on an EC funded research and development project, was selected for testing because it is representative of the class of information providing web site and we had direct access to its designer. Internet experienced web designers and users were selected as the participants. The web site designer subsequently assessed the results obtained by these methods in terms of their usefulness for design purposes.

The main findings of the study were:

1. Each method to evaluation contributes in specific and complimentary respects.
2. Triangulation of method outcomes enhances problem understanding.
3. Some WAMI problem were not relevant to the UNITE site.
4. Some WAMI problem statements were too abstract to directly inform design thinking.
5. Additional to conventional HCI issues, the designer indicated the following as web site design usability goals:
    i. Engagement (keeping the customers at the site);
    ii. Loyalty (engendering customer commitment, thereby motivating customers to revisit the site).

A survey of user and designers needs and expectations

On the basis of the literature review and the above study three main issues were identified for further research:

1. Websites are different from conventional software
   The web designers' expectations of a web site revealed additional goals, i.e., to LEMES, such as that users should like a site, should keep returning to a site, should increase in number and should find what they want.
2. Designers and users probably have different views about web site usability

From the literature, it appears that the web designers perform web site tasks better than non-web site designers, e.g., they learnt faster and their error rate was lower. This suggests that designers have specific knowledge that allows them to predict how the information in a site is organised. As Norman (1990) notes, "The designer is not a typical user". Additionally, designers have more intimate knowledge and understanding of the systems they are developing than the average user will ever acquire (Microsoft Corporation, 2000). Such differences in perspective may lead web site designers to design sites that fail to match users' needs.

3. Problems need to be clearly defined
   The outcome of a usability evaluation should be a clearly defined report in which problems are presented systematically and in a manner that designers can use. In our study, ill-defined problems were related mostly to users' subjective opinions about the system. For example, the statement, "the web site is not attractive", says nothing about why it was judged unattractive or how it might be made attractive.

These matters were explored further through an Internet questionnaire directed at both users and designers. Designers will be asked to indicate what they really needed from a web usability evaluation, e.g., how results should be presented. Additionally, the questionnaire sought to clarify differences between web designers' and normal users' understanding of web site usability. In total 65 web designers and 69 experienced web users responded, ranging in age from 16 to 34 years.

The results of the questionnaire confirmed that designers thought issues such as that users should like a site, should keep returning to a site, should increase in number and should find what they want were important and should be evaluated. Analysis of questionnaire responses enabled each web site design goal to be related to the factors affecting user satisfaction and the web site components influencing those factors, Figure 1.

Designers goals	User issues	Web site features
1. High quality HCI	Easy to navigate, easy to find the information, easy to understand the site structure	Navigation
2. User retention	Download speed, ease of use, attractiveness, layout, colour scheme, helpful information, site updates	Navigation, visual, information
3. Providing appropriate user services	Easy to use and understand structure, helpful information, layout and updates	Navigation, information
4. Likeability	Image, layout, useful information and functions	Visual, functionality, information

Figure 1. The relations between designers' goals, user issues and web site components

Furthermore, some differences in the perception of usability were identified which may affect the manner in which designers amend a site. For example,

we found that users expected a web site to provide diverse functions and information, whereas the designers sought to design a web site around a specific theme, and users put greater emphasis on information content than visual design, whereas these priorities were reversed for the designers.

Regarding the usefulness of current methods, about 50% of the web designers suggested that seeing users using their site is the most useful way of understanding user problems. They felt that the problem statements arising from the usability questionnaire needed a higher degree of specificity to help them redesign and they also required information on users' routes to task completion and whether the site was satisfactory for its target audience.

In summary, the questionnaire study affirmed the issues that had emerged from the earlier research. It established that web site designers seek to attain goals (i.e., 2, 3 and 4 in Figure 1) that are not accommodated within typical HCI or web site evaluation methods. Similarly, there are additional web site user needs that should be evaluated (e.g., a web site that changes over time). It established that designers and users hold different views about web site usability. Clearly, web site evaluation methods should give due emphasis to user needs and should accommodate the additional user needs associated with web site usage. Finally, the questionnaire study confirmed that designers are not content with the typical problem identification reports produced by evaluators. Essentially, these reports need to identify more clearly the nature of a problem, when and how it occurred and the user attitude to it.

Development of a new approach to web site evaluation

On the basis of the studies described above a new evaluation method has been devised. The evaluation techniques employed in the initial study have been extended and expanded in order to accommodate the additional designer goals and user needs derived primarily from the questionnaire survey. The method employs an adapted form of MIMA, card sorting, an adapted version of the observational strategy used in the initial study and a questionnaire tailorable to the web site design goals of the particular site under evaluation. Figure 2 shows each design goal related to example user needs and the methods used to evaluate them so as to assess design-goal attainment.

Designers goals	Examples of user issues	Proposed method(s)
1.High quality HCI	Easy to navigate, easy to the information, easy to understand the site structure	Observation (verbal protocol) route taken, time and MIMA, card sorting
2.User retention	Ease of use, attractiveness, layout, colour scheme, helpful information, site updates	MIMA, card sorting, tailorable questionnaire, observation
3.Providing appropriate user services	Easy to use and structure, helpful information, layout and updates	MIMA, card sorting, tailorable questionnaire, observation
4.Likeability	Image, layout, useful information and functions	Tailorable questionnaire

Figure 2. Methods related to user needs and design goals

Rather than taking the 'one size fits all' approach to evaluation, we believe that the starting place for evaluation has to be the designer. S/he sets the agenda for the evaluation in terms of usability goals (e.g., in terms of time taken to complete tasks and failure rates, etc.), the prioritisation of web site components (i.e., information, navigation, visual and functionality), and the tasks that users will need to carry out. Hence, the evaluation process begins with an examination of the designers' expectations of their system, Figure 1. This enables the results to be presented in terms of the designers goals, which is expected to make the results more meaningful to them (i.e., the final stage in Figure 3). Representative sample users are then subjected to in depth testing using the methods outlined above in the order of Figure 3.

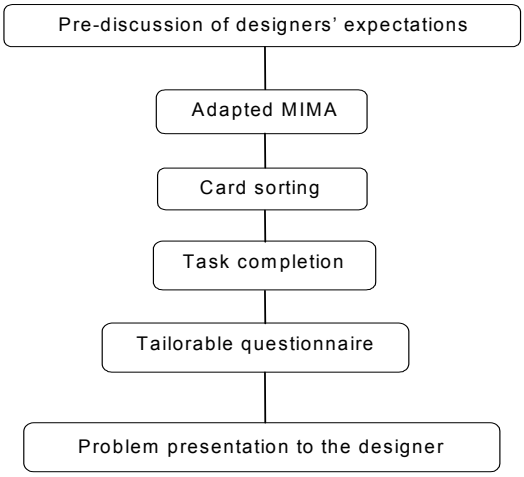

Figure 3: The stages in the evaluation process

Evaluating the method

Having developed the method on the basis of the results emerging from the two studies described above, our next task was to test the method in order to establish whether or not it contributed to the third project objective, i.e., to overcome the limitations of current methods. The aims of this evaluative study were to test:

- Whether the results provided information relevant to both the users' and designers' needs (as identified in the Internet survey) such as for clear page layout and helpful information, etc.
- The applicability of the evaluation results, i.e., whether the results were understandable and useful to designers when redesigning the site.

Superb-Call.com was selected (with the providers' support) as the site to be used in the trial. Many phone cards are available to consumers that differ in terms of call tariffs. This site enables users to compare cards and buy online those best suited to their needs. The target users, as defined by the design team, are overseas students and International business people located in the UK and possessing basic competence in reading English. The site design

team is composed of a marketing manager (with overall responsibility for the site) an interface designer and a computer programmer.

This evaluation was conducted in laboratory conditions enabling network factors, such as download speed, to be controlled. The tasks to be used in the test were developed through discussion with the site designers. Three main tasks were selected: 1) finding out the cheapest phone card using a rate checker; 2) registering as a site member and completing a membership survey; 3) selecting a phone card and paying for it. The method requires a design team to provide task completion time criteria so as to enable the severity of any task completion problems to be judged and in this case they were respectively 2, 5 and 5 minutes.

Six female and four male International students, aged between 16 and 34 years and with higher than intermediate level English, participated in the evaluation. The participants had all visited more than 51 sites and thus had basic competence in using Internet and the site browser (i.e., Microsoft Internet Explore 5.0).

Results

Coverage of web site usability factors (Aims 1 above)

The method returned results relating to:

- Conventional aspects of HCI usability, i.e., attractiveness, controllability, efficiency, helpfulness, learnability and satisfaction. For example, the site was not especially liked and was barely regarded as usable by users.
- Navigation, e.g., most users misinterpreted specific terms and the sorting task revealed a number of links with high miss classification rates, i.e., of a link to site the relevant site information pages.
- Task completion – observation revealed task times both within and in excess of desired task completion times set for the site by the design team. The method also discriminated effectively between tasks satisfying design criteria and those characterised by task completion failure and error.
- Information – again the method effectively discriminated between information that met user and designer needs and that which did not.
- Visual – similarly, the method provided information on those visual aspects of the site that satisfied designers' expectations and those where the users' level of satisfaction was problematic.
- Functionality – finally, the methods revealed which functions were, or were not, valued by users or problematic to them.

Overall, the method was judged to perform well in its power to provide information relating to system performance both in terms of conventional HCI factors and the additional web site related factors established through our prior work.

The appropriateness of the information obtained for design purposes

To assess the practicability of the methods, the problems were reported to and their value discussed with the design team, which was also invited to consider possible redesign solutions.

Following the problem report presentation the designer team agreed that a total of seventeen significant changes to the design were needed. Furthermore, they confirmed verbally that most of the problems were intelligible. For each of the proposed design changes the team offered precise solution ideas which both revealed understanding of the problems and that the problem information was adequate to determine the nature of a problem and when and why it had occurred. The team found the severity rating helpful, but reported that in practice they were unlikely to remedy all problems. Finally, the designers confirmed that the relating of problems to web site components, i.e., navigation, visual, information, and functional components encouraged precise and comprehensive consideration of web site problems and solutions.

Overall the results support the claims made in regard to the contribution made by the method to web site evaluation, both in terms of addressing usability factors specific to this mode of human-computer system and the intelligibility and relevance of the information obtained for design purposes.

Conclusions

We have argued that using a web site design is not the same as using a typical PC application and therefore that the web site usability goals/needs of providers and users may not be the same. We have also argued that the web site user's and designer's perceptions of what makes for a usable and effective web site may differ. Finally, we have argued that current evaluation methods produce problem reports that are less intelligible and useful than they could be.

We have described research that has provided evidence in support of these propositions, e.g.:

- Of design goals, such as that users should like a site, should keep returning to a site, should increase in number and should find what they want.
- Of difference in perspective, such as users expecting a web site to provide diverse functions and information, whilst designers seek to design a web site around a specific theme.
- Of additional web site user needs, such as change in web site content over time.
- Of lack of satisfaction with problem reports, such as inadequate specificity for design purposes.

We have also described the design and testing of a method designed to address current limitations, thereby contributing to web site evaluation. The results of the evaluation are sufficiently encouraging to warrant revision of the method, evaluation of it in different Internet contexts and its formalisation to enable other evaluators to use it.

Acknowledgement

The work presented in this project was conducted with the cooperation of EntaMedia Ltd. Thanks are due to Jason Tsai, Ryan Lee, David Lin, and Margot Lee for their support.

References

Berkun, S. (2001) The Role of Flow In Web Design. Microsoft Corporation, January/February 2001. http://msdn.microsoft.com/library/en-us/dnhfact/html/hfactor10_1.asp?frame=true, 24/11/2001

Bourges-Waldegg, P. (1998) Handling Cultural Factors In Human-Computer Interaction. Unpublished PhD, University of Derby, U.K.

Bourges-Waldegg, P. & Scrivener, S.A.R. (1998) Meaning, The Central Issue in Cross-Cultural HCI Design. Interacting with Computers, 9, No 3, pp. 287-310.

Bourges-Waldegg P. & Scrivener S.A.R. (2000) Applying and Testing an Approach to Design for Culturally Diverse User Groups. Interacting with Computers, 13, 2, pp. 111-126.

Brinck T., Gergle D., & Wood S. D. (2002) Usability for the web, Morgan Kaufmann Publishers, USA

Gould, J.D. & Lewis, C (1985). Designing for usability--key principles and what designers think. Communications of the ACM, 28, 3, pp. 300-311

Kirakowski, J. and Cierlik, B. (1998) Human Centered Measures of Success in Web Site Design. Fifth Human Factors and the Web meeting, June 1998.

Nielsen, J. (1992) Evaluating the Thinking Aloud Technique for Use by Computer Scientists. In: Advances in Human-Computer Interaction, Vol. 3 (eds.) H. R.Hartson and D. Hix, Norwood, NJ: Ablex. pp. 69-82.

Nielsen, J. (1993) Usability Engineering. Boston: Academic Press, Inc.

Nielsen, J. (2000) Designing Web Usability: The Practice of Simplicity. Indiana, USA: New Riders Publishing.

Norman, D. A. (1990) Why Interfaces Don't Work. In: The Art of Human-Computer Interface Design (ed.) B. Laurel. Harlow, England: Addison Wesley Longman Inc.

Preece J. (1994) Human-Computer Interaction. Harlow, England: Addison-Wesley.

Rubin J. (1994) Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. USA: John Wiley & Sons Inc.

Spool, J. M. (1999) Web Site Usability, A Designer's Guide. USA: Academic Press.

Urquijo, S.P., Scrivener, S.A.R. & Palmen, H. (1993) The Use of Breakdown Analysis. In Synchronous CSCW System Design. Proc. of the Third European Conference on Computer-Supported Cooperated Work. Milan: Kluwer Academic Publishers.pp. 281-293.

Usability in Software Design. Microsoft Corporation, October 2000. http://msdn.microsoft.com/library/en-us/dnwui/html/uidesign.asp?frame=true, 24/11/2001