Website Designer as an Evaluator: A Formative Evaluation Method for Website Interface Development

Chao-Yang Yang

College of Management, Industrial Design Department, Chang Gung University,
259 Wen-Hwa 1st Road, Kwei-Shan, Tao-Yuan, Taiwan, R.O.C.
dillon.yang@mail.cgu.edu.tw

Abstract. Commerce plays a fundamental part in a lot of websites so that their goals may be different from conventional computer system design e.g. to increase the user base or encourage repeat visits. With limited budgets, website designers are unlikely to involve their users during the design process and not all website designers have access to an evaluator, appropriate testing facilities or evaluation knowledge to support their design. The research develops a low cost, tailorable, formative evaluation method for web designers. The method addressed both HCI and commercial website goals such as the encouragement of repeat visits. This research first investigate the contemporary evaluation method, the users' and designers' needs from websites and website evaluation methods. Finally, the method was developed as a set of guidelines and verified in the evaluation of a website. The potential usefulness, practicality and necessity of the method was then confirmed by website.

Keywords: Website usability, Engagement, Formative Evaluation.

1 Introduction

Instead of evaluating websites with real users, designers normally rely on their own experience, company guidelines, and clients' demands to inspect the site before launch. Such an approach yields user responses such as "I don't like the colour", "I cannot find the information I want", and "I don't want to provide that much information when I shop online". For designers to fully realize a site's market potential, an appropriate website evaluation method is needed that will enable them to understand user requirements and enable them to redesign a site in terms of user needs.

The development of website technology such as Java, My SQL database system, has made the development of various functions for users possible. However, such innovations complicate the design task by increasing the number of elements from which a designer can select to include in a website. As website and software interfaces are similar in some respects (such as information architecture and navigation design), a number of standard HCI evaluation methods, such as observations and questionnaires, have been used to help web designers identify and categorize usability problems in terms of effectiveness, efficiency, learnability, and satisfaction. HCI is concerned with understanding how people use computer systems with the aim of informing the development of systems that more closely meet users' needs [1, 17].

J.A. Jacko (Ed.): Human-Computer Interaction, Part I, HCII 2009, LNCS 5610, pp. 372–381, 2009.

Attracting new users and retaining them through good design and usability are of greater importance as competition increases on the Internet. However, because of the inherent differences between website usability and conventional computer systems usability, it may be inappropriate to directly apply standard HCI methods to evaluate websites. Spool et al [13] and Nielsen [9] suggest that website designers should pay more attention to enhancing the functions and information to make users like a web site. Furthermore, Nielsen [9] has also established that users have low tolerance for complex designs or slow sites; people don't want to wait and they don't want to learn how to use a home page. As there is no training or manuals for a Web site, people have to be able to grasp the function of the site immediately upon scanning the home page. On the other hand, this research has identified a need for designers to also consider the site's ability to retain users and attract regular repeat visits. Hence, the users' needs from the website and the designers' needs from website evaluation should be clarified.

As websites develop and evolve more quickly and cheaply than normal software releases [3, 16], designers are commonly faced with the need to refine or redesign their sites. Testing late in the development cycle enables the site to be compared to predetermined usability standards or benchmarks (e.g. task components work together, thereby preventing flawed releases entering the marketplace, that will need recall or adjustment [12]. Thus a general method that can be used quickly, by practicing designers is needed that will provide the specific information needed for redesign.

Most existing evaluation methods have been criticized because they do not specifically identify usability problems [2]. Such deficiencies in information about user issues may leave designers guessing at solutions. For example, the statement "some of the users cannot find the correct navigation to link to next page to finish the task" provides insufficient information for an evaluator to judge the problem precisely. The best that can be concluded is that "the users cannot navigate correctly", which is not specific enough to properly guide redesign. Instead, the actual problem could be that the users do not understand the navigation term used, or that the navigation element could not be located easily. When a problem is not fully understood [10] or is described at the wrong level of abstraction, and the designer is not a typical user [11], it is easy for the designer to overlook some of the most critical but subtle dimensions that contribute to a situation, and the resulting solution may make some parts of the interface worse [10]. This being the general case, it is important to examine current evaluation methods to establish their specific deficiencies.

When the above in mind, this research aims to contribute to knowledge by enhancing existing evaluation methods forward in website design. A low-cost evaluation method for web designers to use for formative evaluation, prior to site launch was developed and validated. In particular the research shows that:

In terms of designers' goals and users' expectations, web design needs to consider more than just the HCI issues considered in conventional computer interface design;

- User-centred design is important, as designers' and users' perceptions of websites are different;
- Web designer's requirements with regard to evaluation methods are not met. They need to be presented with well specified, detailed problems from which they can

generate effective solutions. Such methods would enhance the efficiency of the re-design process;
- Decomposing the website into visual, informational, navigational, and functional sections enables the evaluator to systematically test and determine problems.

2 Methods

The research method adopted was that of problem-solving. Given that a problem has been identified, requirements for the solution generation are collected; a solution is proposed and finally tested. In order to achieve our aims and objectives, a number of methods were used as appropriate to an "understand –propose –realise -evaluate" life-cycle [17] (as shown in Fig. 1). An understanding of the problem was achieved through a literature review, analysis of current methods and attitudinal survey. From the requirements identified in the survey, literature review and the analysis of the applicability of existing evaluation methods, a web evaluation framework was developed that would meet both users' and designers' needs. An action learning approach was taken to the development of the method, whereby the researcher iteratively designed, tested, and selected methods based on their usefulness. Following evaluation (see below), the method was formalized for use by practicing designers. The method was subjected to three forms of evaluation: Iterative interface design and evaluation; formalisation of the method; evaluation by web designers.

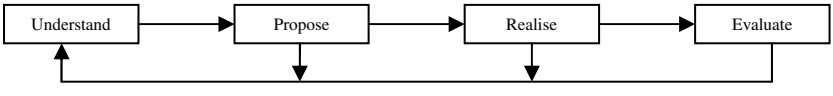

Fig. 1. Design Research Model with Feedback Loops [17]

3 Literature Review

A website consists of navigation, information, and visual elements. The purpose of a site may be viewed as a conjunction of satisfying what users are trying to accomplish (e.g. doing research, buying products, or downloading software), and the designer's and client's goals [13]. These goals are related to issues such as HCI, pleasure, security, technical issues, and accessibility.

Typically, the website designer is in charge of the look, ease of use, and the content of a website [3], with the intention of providing a clear marketing message, trust, frequent update of information, aesthetics, functionality, so as to achieve client company goals. The following elements have been identified as important in terms of the user and designer's needs.

1. *Navigation.* HCI plays an important role in website navigation design. This has been addressed in website design in terms of effective and efficient information structure, user interface, page design, content authoring, cognitive process, linking strategy and task design.

2. *Information.* A commercial website should provide useful, helpful (e.g. to help make purchase decision), updated, and individualized information. In addition, it should include clear marketing messages and effective privacy statements that show it is following privacy and consumer protection guidelines, making the security of customer data a priority and using independent certification bodies.
3. *Visual design.* The likeability and attractiveness of visual design relies on the aesthetics of the layout, colour scheme, animation etc. Also accessibility for colour blind users needs to be considered.
4. *Other issues* – in particular technical and accessibility issues should be considered.

This section has introduced website design, considered the role of HCI and identified a new set of web design considerations. Evaluation has been identified as crucial in supporting the design process, hence it follows that website evaluation methods will need to accommodate the identified design goals and issues associated with website design. The diversity of website users, their purpose, individual characteristics, information seeking behaviour, and cultural issues are complicated and affect the design of different types of websites.

Having reached this point the question is, "To what extent do website design and evaluation methods handle the factors discussed above?" With this in mind, next, website design methodologies will be reviewed prior to an examination of website evaluation methods. The goal of this analysis will be to assess the extent to which current website design and evaluation is fit for purpose.

Effective usability evaluation methods for website redesign need to be fast, low-cost, easy to learn, provide high confidence and high impact. This section reviewed conventional usability methods and those that have been employed in website design. It can be concluded that most website usability evaluation in late development stages requires real user testing, i.e., observing users completing website tasks. These methods provide reliable and useful information for redesign.

However, given that existing methods have been adapted from those used in HCI, there are several shortcomings in the extent to which website features and issues can be evaluated using such methods and the extent to which such methods can be applied to rapid website development with short product life cycles, small development teams and limited budget. In such an environment, methods that require specialists or specific equipment may not be used. Further, the evaluation methods have also been discussed from a marketing perspective in which it has been shown that the marketing goals have not been properly addressed. In the next section, the issues identified from this review will be investigated through a study of selected user testing and data analysis tools.

4 Study of Current Usability Testing Methods

To gain more insight into the usefulness of current evaluation methods, a representative set of different types of methods was used to evaluate a website prior to launch. These included observation, Meaning in Mediated Action [15], and Website Analysis and Measurement Inventory (WAMMI) [5] and Breakdown Analysis [14, 18]. The UNITE (Ubiquitious and Integrated Teamwork Environment) website, which had not

been launched was selected as the test object. This site was designed to promote an EU founded project which aimed to develop an environment for virtual teamwork.

Overall, the evaluation was time consuming (especially in the task completion section) and each method provided both useful and not so useful information for redesign. By observing the process and the usefulness of the results, WAMMI was identified as being useful in assessing the participants' preferences through its rating system. However, some of the questions were irrelevant and unclearly defined problems were not useful. The designer indicated that the MIMA and Task Completion sections were helpful for redesign of the navigational elements as they provided details of specific navigation problems. Further, the designer indicated that WAMMI provided information about the site's engagement and time-based issues which was lacking in MIMA and the User Testing. To summarize, each method had strengths and weaknesses and using more then one technique can help ensure that the findings are reliable [14]. The comprehensive information needed for redesigning the site can be generated through multiple methods, although the process could be shortened by discarding less useful elements, simplifying them and concentrating on the elements/information needed by the designer for the task in hand.

In summary, the results from existing different methods support each other; the results from user testing and MIMA are useful for redesign; marketing issues need to be considered in the evaluation; marketing goals can be addressed in a questionnaire but this should provide more specific information. The next stage of the research will employ a questionnaire to gather further opinions about website usability from users and designers. This will help to establish the requirements for website evaluation enabling us to construct a new method, geared to the needs of designers.

5 Internet Surveys of Designer and User Needs

Through testing existing usability methods on the UNITE website it was found that different methods favour different aspects of web usability. Through triangulation and selected use and adaptation of different methods a more complete picture of usability issues can be established. However to be useful outside the experimental situation such a combination of methods has to provide sufficient detailed information for designers to concentrate on the important elements from the user's perspective. Therefore the design of an effective method should, on the one hand recognize the needs of the designers (to produce usable sites quickly) and the requirements of the design task and on the other hand the needs of the user – to find the information they need efficiently.

Taking the findings of the previous studies into account, the important elements for web site design may be summarized as: adherence to best practice in HCI, usefulness, pleasure in using the site, user retention, and the ability to attract new users. This part of research will detail a study undertaken to establish whether there are any differences in the way in which web site designers and users perceive usability, and what type of information designers would like for redesign (i.e. formative as opposed to summative evaluation).

In this case the method chosen was an on-line Internet questionnaire which would help reach the massive key target populations – web designers and site visitors. This

questionnaire focused on the users' and web designers' opinions of web usability. Web designers, whose sites were included on www.coolhomepages.com, were invited. The user participants were selected by posting an invitation on professional message boards and discussion areas (for experienced users) such as www.coolhomepages.com and www.msn.com, as it was believed these users would be more web-savvy.

The results have confirmed that all five general goals should be given equal prominence in website design. In addition, several participants felt that a website should require its users input (e-mail or buying products) as this can bring benefits to the website. These can be attained by improving design requirements such as ease of navigation, helpful information, and good visual design. Helpful, updated, or interesting information are the user's primary needs from a website. These features also affect the likeability of and degree of user engagement with a site. Clear and attractive visual design mainly affects likeability. Easy of navigation is a primary requirement, a feature emphasized in conventional HCI. In addition, as described previously, ease of navigation has been indicated as an important predictor of recommendability. Functionalities, such as a message board and search engine have been indicated as key ingredients of a website. Therefore, it is reasonable to propose that providing useful functionality that meets users needs, could improve the degree of engagement with and likeability of a website.

The designers generally pay attention to the site's usability before launching the site. Although the designers stated that they understood typical usability statements and could act on them appropriately – when their answers were considered in detail, different solutions were proposed by different designers to the same problem, indicating that the statements might be ambiguous and lacking in sufficient detail for reliable decisions to be made. Designers prefer feedback from users that have a clearly stated problem report relating to the site's information structure, image, colour, compatibility, font, symbols, logo, etc. Further, they were concerned about the cost and complexity of evaluation. However, given the differences between designers and users it is still necessary to evaluate websites with real users.

To summarize, a good site is designed considering issues of adherence to best practice in HCI, usefulness, user retention, likeability, and ability to attract new users. Designers and users were shown to have slightly different views on website usability. In terms of evaluation, the existing problem statements were not detailed enough.

Following Newman and Landay's [7] categorization, websites can be broken down into navigation, information, and visual elements to provide a clear view of the site. In addition to these elements, functionality has also been identified as a fundamental element that may affect website goals and usability.Applying this categorization to the development of an evaluation method may provide a clearer view of the website and lead to the development of more designer friendly methods.

Collecting data is necessary but not sufficient for a usability test [4]. A clearly defined problem report is also necessary and the need for a usable output should be remembered, for example detailed problem identification is required to avoid ambiguity, and questions should be related to the actual web site rather than overall features. Such problems may arise especially in areas of overlap between knowledge domains. For example a problem such as "this web site is a waste of time in every respect" could relate to either poor quality information or hardness of navigation. To avoid

such ambiguity the method will relate statements more clearly to the domain they refer to, such as aesthetics or navigation.

The results of this study have also shown that depending on the site, the target users may be different and may have different perceptions using the site. For example, a human resources site may aim to provide an easy to use interface, but a Disney site may aim to achieve high likeability. Without feeding the target user's needs, the site may fail to retain their users. Therefore, a more tailorable approach to evaluation is needed, which is based around the site, the expectations of the owners of the site, the designers and the users.

6 Composing the Website Evaluation Method

Previous research has shown that increasing the user base, the likeability and efficiency of the web site, engaging the users, and identifying and meeting user needs are important for website development. For users, as shown in Fig. 2, helpfulness of information, ease of use, attractiveness of the visual design, and functionality play a part in determining whether a site achieves these goals. An evaluation method is proposed that can be used by designers. The method is composed mainly by four evaluation techniques: MIMA interview, card sorting, user testing and structured interview. It takes into account time, cost, learning time, degree of confidence in the method, and the potential for impact on redesign. The evaluation method was designed to provide an effective and efficient formative evaluation that could be used by designers to provide information for redesign.

MIMA interview. The elements to be tested are shown to the participant first in isolation, and then in the context of the web page they appear on. The participant is required to interpret the representation. This is recorded in the format of Table 1 against that of the designer. Where necessary the evaluator should ask for clarification of interpretation, so that the nature of the interpretation is fully understood.

Table 1. An example of function key assessment

Functions	Intended action	Assessment of IM	Participant's interpretation
Search	Start searching the given keyword in the database		Search information related to the keyword

Card sorting. The participant is asked to associate cards with the most relevant main navigational links (could be terminology or graphic). The cards containing the navigational elements to be tested are placed separately on the desk in front of the participant with the main navigation cards set out at the top. The participant needs to assign the sub-navigational elements or contents to these by placing them underneath the element to which they appear most relevant. If the site contains sub-sub-navigations/ contents, the evaluator should then ask the participant to assign these under the sub-navigations determined by the designer.

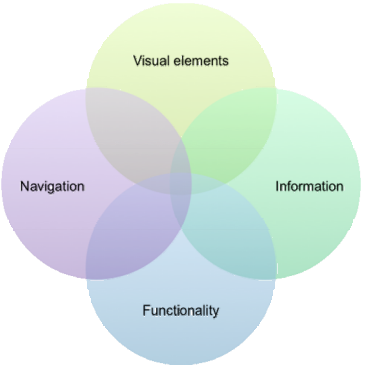

Fig. 2. Relationship among the design elements

Table 2. The card sorting result format

Main navigation	Designer's categorization	Assessment of participant's categorization	Participant's categorization
Superb-cards	"Saver" telephone card	◯	
	"Bubble" telephone card	-	
		+	30% off Swifty telephone card
		+	Advantage buying from Superb Call
	download mobile ring tones and pictures	-	
(yahoo.com.tw, 2005)		+	Web-telephone

The participant's categorization is recorded and assessed as shown in Table 2. Those instances where the participant categorizes information in a similar way to the designer are of no interest and are marked "◯". Additionally, a "-" is given when the participant fails to sort a card in the correct place on the navigational link. A "+" is given when a card is placed on an incorrect link. This information may assist the designer in re-organizing the navigational structure.

User testing. Direct observation provides more objective information than surveys [6]. The user should now have some familiarity with the site, and is provided with a set of tasks to assess how efficient the web site is in letting them perform common tasks, where problems occur and the reasons for these. The tasks are fully explained to the users, but no other assistance is provided. The participants are required to verbalize their thoughts and feelings during the task as this can generate valuable usability information [8]. The time, path, action, and verbalizations are recorded as in Table 3.

Table 3. The task completion data analysis format

Time (second)	Path	Actions	Think aloud protocol
5	Home	Moving the cursor around all links in this page	Still looking, still looking. Haven't seen any thing say "subscribe" at the moment.
22	News		I am going to News area as it looks most related

The path and completion times will later be compared to those provided by the designer. Where errors occur, the verbalizations and video are used to provide a rationale for this.

Structured interview. The structured interview is conducted to assess information, visual, and function design. The interview is structured around a questionnaire, with the participant being required to provide a rationale for their ratings. As the questions are closely aligned to the contents of the web site, this necessitates participants using the website in some detail. All the ratings, for each question are combined for all the participants and average scores are used to determine the severity of the problems. Examples of the rationale are also presented so the designer can achieve a greater understanding of the design problem.

7 Conclusions

This research considers the requirements to support commercial website design based on user's and designer's needs. Typically, HCI plays an important role in this domain and recent research shows that websites require more aspects to be taken into account. Without addressing specific issues such as marketing and pleasure, current usability methods will poorly support the design. Hence, the appropriateness of applying standard usability measures to website design was investigated.

By incorporating the user's and designer's opinions, it was confirmed that websites not only have to meet usability criteria, they also have to increase the user base, likeability etc. It also showed that these issues can be achieved through improvements to the design components in navigation, information, visual, and functional aspects. Each aspect can be assessed efficiently and precisely by different evaluation techniques. Therefore, a multi-method method has been produced which is tailorable to different websites to advance the use of existing usability evaluation in commercial website design. In addition, the research has formalized the method into one which a designer can use. The studies undertaken have shown the validity, practicability and usefulness of this approach for website designers.

In conclusion, the research has contributed to knowledge by identifying and filling the gap in the current use of evaluation methods by providing a method that practicing web designers can use with representative end users.

References

1. Benyon, D., Davies, G., Keller, L., Preece, J., Rogers, Y.: A Guide to Usablity. UK, The Open University
2. Berkun, S.: The role of flow in web design, Microsoft Corporation (1990), http://msdn.microsoft.com/library/en-us/dnhfact.html/hfactor10_1.asp?frame=true (accessed 24/11/2001)
3. Brinck, T., Gergle, D., Wood, S.D.: Designing Web Sites that Work - Usability for the Web. Academic Press, USA (2002)
4. Dumas, J.S., Redish, J.C.: A Practical Guide to Usability Testing, Intellect (1999)
5. Kirakowski, J.C.N., Whitehand, R.: Human Centered Measures of Success in Web Site Design. In: 4th Conference on Human Factors & the Web, New Jersey, USA (1998), http://www.research.att.com/conf/hfweb/proceedings/kirakowski/ (accessed 06/04/2000)
6. Moseley, B.: Test the Usability of Your Web Site. Folio:PLUS 30(Part 4), 9–10 (2001)
7. Newman, M., Landay, J.A.: Sitemaps, Storyboards, and Specifications: A Sketch of Web site design Practice. In: DIS 2000, New York (2000)
8. Nielsen, J.: Usability Evaluation and Inspection Methods. Addison-Wesley, Reading (1993)
9. Nielsen, J.: Designing web usability: the practice of simplicity. New Riders Publishing, USA (2000)
10. Nielsen, J., Mack, R.L. (eds.): Usability Inspection Methods. John Wiley & Sons, New York (1994)
11. Norman, D.A.: The Design of Everyday Things. New York, Doubleday (1990)
12. Rubin, J.: Handbook of Usability Testing: how to plan, design, and conduct effective tests. John Wiley & Sons, USA (1994)
13. Spool, J.M., Scanlon, T., Snyder, C., Schroeder, W., DeAngelo, T., et al.: Web Site Usability: A Designer's Guide. Academic Press, San Diego (1999)
14. Urquijo, S.P., Scrivener, S.A.R., Palmen, H.: The Use of Breakdown Analysis in Synchronous CSCW System Design. In: Proc. of the Third European conference on Computer-supported Cooperated Work, Milan, Italy (1993)
15. Waldegg, P.B.: Handing Cultural Factors in Human Computer Interaction. Derby, UK, doctoral thesis (unpublished, 1998)
16. Winckler, M., Pimenta, M., Plalanque, P., Farenc, C.: Usability Evaluation Methods: What is still missing for the Web? In: 8th International Conference on HCI International, New Orleans, USA, August 5-10 (2001)
17. Woodcock, A.: Supporting Ergonomics in Concept Design, Loughborough, Loughborough University, doctoral thesis (unpublished, 2001)
18. Woodcock, A., Scrivener, S.A.R.: Breakdown Analysis. In: McCabe, P. (ed.) Conference of Contemporary Ergonomics, pp. 271–276. Taylor and Francis, Edinburgh, UK (2003)