To measure or not to measure:

Why web usability is different

from traditional usability.

Charlotte Olsson
Department of Informatics
Umeå University, 901 87 UMEÅ, Sweden
+46 90 786 6820
colsson@informatik.umu.se

Abstract: Web usability is a common term used in discussions of WWW (World Wide Web). This is definitely important, as more and more web sites are frequently visited and have great impact in the new economy, the network economy.

However, research lack in consistency in defining 'web usability'. Some point out that the term is equivalent with the traditional term in HCI (Human Computer Interaction) research. Others point out the differences with traditional usability and do this with comparisons between the two.

This article discusses the concept of usability in general as well as the notion of web usability and their relation. Further, a conceptual model is built to facilitate how to handle usability aspects on the web. This model is then exemplified with two examples of web sites. Here, the importance of choosing the right categorization is clear. The sites are both e-commerce sites, however, they show upon clear differences when it comes to usability aspects.

Web usability is an important issue when it comes to design of web sites. We have to start to take it seriously.

Key words

Usability, web usability, world wide web, www, the web, e-commerce, learning environments

Introduction

As the World Wide Web (www or just 'the web') continues to grow, as if out of control, the need for a discussion around usability aspects on the web emerges. So far, in relation to the number of web sites, the usability aspects more or less have been left out. One explanation of this phenomenon may be the fact that web designers are working as quickly as possible. They do so in order to build as many web sites as possible because of the overheated demand and therefore do not have time, or simply do not need, to usability test their designs. This may though be a very simple explanation on a more complex problem. The fact is, even if designers on the web wanted to call for backup in usability engineers to do usability testing on their sites, problems occur simply because of the nature of the web. Below, aspects relating to usability on the web are shown:

First, this medium is quite new. We know very little about how to design for hypermedia, as the technology behind the scene is called. New ways of structuring information is needed, and this has to be tested. Right now, users get frustrated when things do not work out the way they are used to (Nielsen, 1999). Standards and guidelines have not yet settled in web design and more, the hypermedia type of technology requires special types of features (Spool et al, 1999). Second, the group of visitors is heterogeneous as the medium is public and this makes the feedback on poor design difficult to reach. For instance, as long as the purpose of the users is diffuse, we do not know what to measure. One site could also have many purposes as selling, reviewing articles, entertainment and more

(Schneiderman, 1997). Third, the group of web designers is also heterogeneous. The reasons for this are the growth of the web as well as the ease of access to facilitating web design. (Bevan, 1998). Design professionals are of course a big group, but in general almost everyone can design web sites. This makes emergence of standards and guidelines difficult (Nielsen, 1999). Finally, the technology is in its nature heterogeneous. Platforms, browser types and versions, html versions and more, make the design a complex issue. The medium was at first mainly intended to be used for academic markup language for distribution of texts in networks. Nowadays, interaction designers create interactive 3D games for the web, and layout is perhaps the most discussed topic in web design. HTML is simply not suitable for this type of usage.(Mayhew, 1998)

People leave web sites all the time because of usability aspects, as they get stuck, and they may never come back. Web usability is different than usability in general, but how do these two concepts relate and how do they differ. That is what this paper discusses.

In order to discuss this usability as a more general concept is first discussed. This is then compared with related work around web usability. What have been said about the concept web usability itself, and what types of studies have been conducted. A framework of how to handle issues of web usability is conducted based on earlier discussions. To test the reliability in the framework, exemplification have been conducted upon some web sites. Finally, this is discussed and further research is exemplified.

The concept of usability in general - some key points

Usability is a key concept in HCI. It is concerned with making systems safe, easy to learn and easy to use (Preece, 1994). The term usability may in daily talk suggest something it is not. Below, a figure showing usability and its context is displayed. Note though that this is one of many categorizations of usability. (For further readings, c.f. Dix et al. 1998)
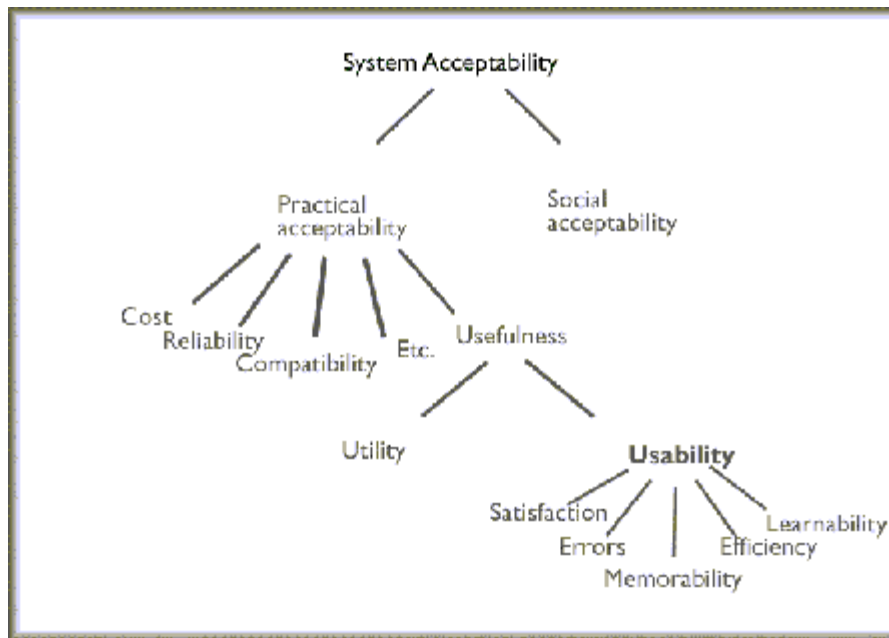

Figure 1: The context of usability in general (Nielsen, 1993, p.25).

In short, descriptions of some of the general concepts above are: (For the interested reader, it is explained at pp. 24-25 in Nielsen (1993).

    System acceptability. Whether the system is good enough to satisfy all need and requirements of all stakeholders, from direct users to customers and more.

    Social acceptability. Whether the system correspond to social rules and norms in the context.

    Practical acceptability. Acceptability according to categories as cost, reliability, compatibility with other systems, usefulness and more.

    Usefulness. Is the issue of whether the system can be used to achieve some desired goal? Can be broken down into utility and usability.

    Utility. A question of whether the functionality of the system in principle can do what is needed.

Usability. A question of how well users can use the above functionality.

Further, Nielsen (1993) defines usability as containing at least the following aspects:
1. Learnability: The system should be easy to learn so that the user can rapidly start getting some work done with the system.
2. Efficiency: The system should be efficient to use, so that ones the user has learned the system, a high level of productivity is possible.
3. Memorability: The system should be easy to remember, so that the causal user is able to return to the system after some period of not having used it, without having to learn everything allover again.
4. Errors: The system should have a low error rate, so that users make few errors during the use of the system, and so that if they do make errors they can easily recover from them. Further catastrophic errors must not occur.
5. Satisfaction: The system should be pleasant to use so that users are subjectively satisfied when using it; they like it.

Usability tests may be conducted in numerous ways, including all from one single technique to a whole repertoire of approaches. It is important to be aware of what to measure. Two common approaches to measuring usability are the following (Redmond-Pyle & Moore, 1995):
Performance tests, where users use the system to perform a task, and their effectiveness are measured. Common measures are speed, accuracy and/or errors.
Attitude surveys, where user satisfaction and user perception of the software is captured. Common ways of capturing data are questionnaires or interviews.

Web Usability - related work
Related work around web usability is divided into two general groups; (1) Methodological papers, where methodological issues related to the concept 'web usability' is discussed, and (2) Reports around results from web site usability tests, with more brief discussions around how to generalize results and more.
Examples of the former are:
Schneiderman (1997) discuss usability aspects related to the web and says that, as in any media, criteria for quality vary with the genre and author's goals. His idea is that there are some web-related criteria that may be seen as more general, like visual appeal, comprehensibility, utility, efficacy and navigability. However, he continues to warn about these high-level goals and point out that a categorization of the web is needed to find more fulfilling criteria to test. The problem though is to find bases to categorize from. Schneiderman gives some examples:
By originator's identity. Individual, group, university, corporation, nonprofit organization or government agency.
By the number of web pages in the site. A similar way is to look upon the amount of information on the site.
By goals of the originators, as interpreted by the designers. Here, the spectrum is wide. From a personal file with chaotic structured information to impressive organizational annual reports. Further, as commercial sites start to grow elegant product catalogs and lively newsletters will be the norm. Web-zines - magazines on the web, digital libraries and much more, all make different kinds of criteria, as well as special usability needs.
By measure of success. For individuals, the measure of success for an on-line resume may be getting a job or making a friend. For many corporate sites, the number of visits measures the publicity. Further, for others, the value lies in the amount of sold articles from the site. Other measure success in diversity in hits or hours spent on site. Example of the latter may be entertainment sites.
Another work done in this research direction is a discussion around how web sites have other kind of characteristics than traditional interfaces (Laskowski & Downey, 1997).
Gaines et al. (1996). Discusses dimensions of problems on the web and try to categorize sites from the concepts of utility and usability. They come up with a layered framework. The article is not further discussed here, though it is interesting work.
Ratner (1998) tries to come up with some conclusions around novice and expert users in learning environments using Netscape. She stresses that even if the goal of the educators, have a specific goal and that the students seem to be a homogenous group, they are not. This must be taken into account in design of such web based learning environments.

Examples of the latter types of related work might also be divided into two subgroups.
So far most usability studies of web sites have focused mainly on efficiency aspects (e.g. the time it takes for a user to find a piece of information in a relatively large site).

Information retrieval is the far most common target for usability testing at web sites. This is because this activity often is seen as central at the web in general (Spool et al, 1999).

Usability tests may be conducted in numerous ways, including all from one single technique to a whole repertoire of approaches. It is important to be aware of what to measure. Two common approaches to measuring usability are the following (Redmond-Pyle & Moore, 1995):

1. Performance tests, where users use the system to perform a task, and their effectiveness are measured. Common measures are speed, accuracy and/or errors.

An example of this is Borges et al (1996)and Borges et al (1998), where they first conduct heuristic evaluation on a number of university sites. Re-design of some of them are then conducted and finally task analysis is done where users are measured when doing tasks. The usability team then ended up with a list of guidelines as a result of their test. However, they are very strict to tell upon the narrow spectrum of web sites these guidelines are appropriate support for design.

2. Attitude surveys, where user satisfaction and user perception of the software is captured. Common ways of capturing data are questionnaires or interviews.

A typical example of this is Spool et al (1999) and their huge usability test of big corporate sites with main focus on e-commerce. This report, or more book, covers the study of nine sites, and here the tests are much more wide. Instead of using the clock in measuring, the test team uses interview forms before and after combined with observations. The users got tasks, but interest was more put on ways of finding information, instead of how quickly the information was retrieved.

Grose et al. (1998) shows with a two folded studie that web style guides differ from traditional style guides and stress the fact that this must be investigated further.

These examples shows how usability engineers handles these aspect in different ways. In the next section, we should create a framework for usability in order to quicker grasp how to conduct tests.

A framework for web usability

In order to come up with the differences between usability in general and web usability, the notation of describing usability in general, is used. This notations is combined with Schneiderman's discussions around categorizing of the web. Below the concept is displayed:
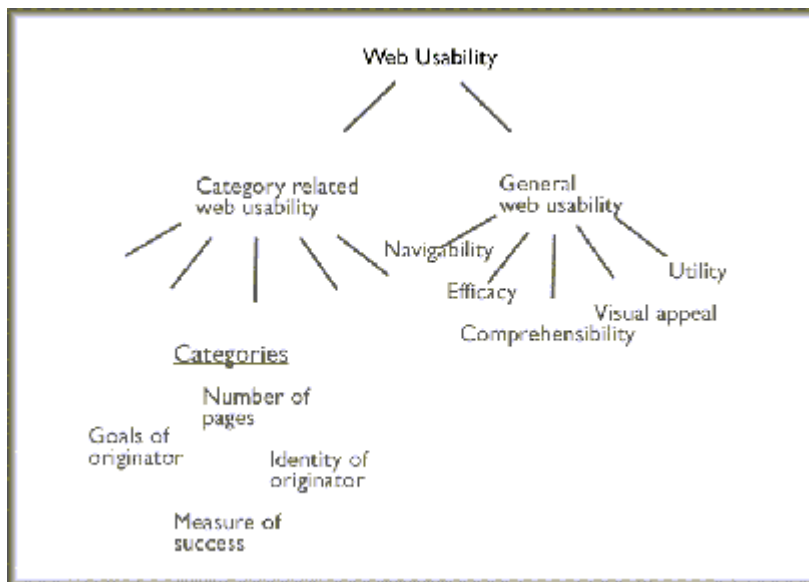

Figure 2: A framework for web usability.

This framework is below tested with two e- commerce web sites, Amazon.com (figure 3a), widely known book store and the 'Robinson' web site (figure 3b), a support site for a famous TV-show at Swedish Television. The two

sites may be categorized similar because of the fact that they both are different kinds of e-commerce sites. However, if other types of categorizations are used, differences start to appear.

For instance, if the measure of success is used, the first, Amazon.com might measure the success in amount of sold items. Another success for them might be amount of hits, as they sell a lot of banners on their site. The Amazon site is difficult to categorize as it is very wide. The company have many purposes with their popular site. Depending on what activity on the site we are designing/evaluating, we must focus on different measures. This, the above figure might help us with.

The 'Robinson' site is quite different. As mentioned, it is a entertainment site, tightly connected with a TV-show. The show is a kind of reality soap opera and on the site you can find supporting information as gossip about the 'actors', screensavers, competitions and more. Also, the design of the site is very different from the Amazon site. On Robinson site there are sound and moving pictures. Overall it gives a very 'Flashy' impression. To test this type of site, time measures are useless. The goal from Swedish Television as well as their measure of success focus more upon high rates of hits, as well as to get the visitor to stay for a while. The visitor on this site want to explore a suprisingly mysterious site and stay for a while. That goal of the designers as well as Swedish Television is quite clear.
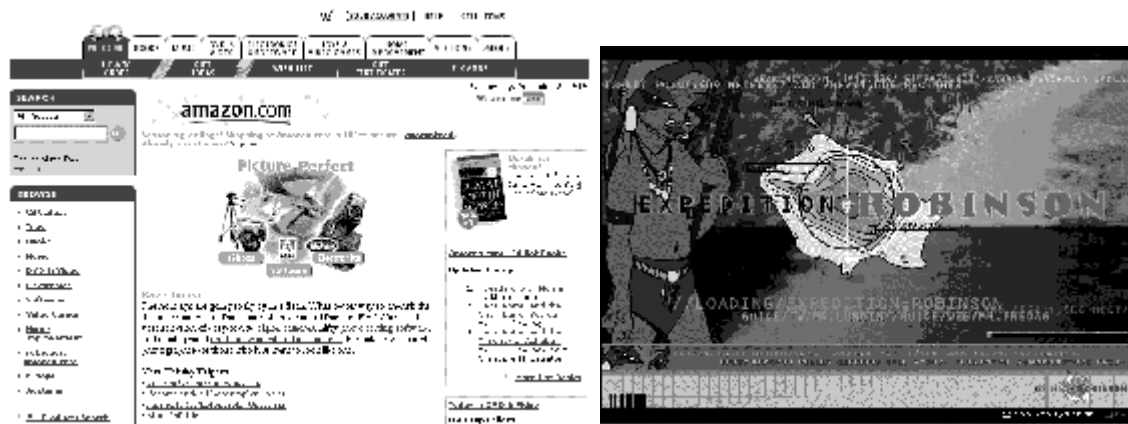

Figure 3: Screenshots of a)Amazon.com and b) Expedition Robinson at the Swedish Television.

Conclusions

Usability on the web is different from usability in general. These conclusions we draw from combining different discussions around the subject and then testing it. The figure of web usability should be used as a guide in order to come up with the proper usability tests and when finding out what the proper measures are. In our cases, with two e-commerce sites, it is clearly shown that they require different measures. This, even if they may be put in the same category of identity, large corporation. In this case a categorization based upon goals of the originators or measure of success may be better. So, what this test and discussion show are:

Web usability is different from usability in general.

Categorization of web sites may be a good way go get/give guidance when discussing web usability.

The way we categorize web sites show strong implication in the way we test usability.

The built model can give some guidance in the work around usability on the web.

Future work

Tests have to be conducted with focus upon the relation between web usability and different categorizations of the web. A research question in this could be; How do the same type of web sites relate in usability aspects and what categorizations are fruitful for this purpose, e.g by genre, by technical aspects or by size of the site. This is interesting because of the need for general guidelines. If, and in that case what, do educational web sites have in common when it comes to usability aspects? Are the technical aspects, as usage of frames for navigation, critical in design for the web? Regarding to usability aspects, how should we categorize the web?

As we do not really know much about these relations yet, this will be important feedback in the debate and research of web usability.

Acknowledgements

References

Bevan, N. (1998). Usability Issues in Web Site Design. In Proceedings of UPA'98. Washington DC.

Borges, J. A., Morales, I., Rodrigues, N. J. (1996) Guidelines for Designing Usable World Wide Web Pages. In Proceedings of ACM Conference on Human Factors in Computing Systems, CHI'96.

Borges, J. A., Morales, I., Rodrigues, N. J. (1998). Page Design Guidelines Developed Through Usability Testing. In Forsythe, C., Grose, E., Ratner, J., Human Factors and Web Development. Lawrence Erlbaum Associates, Publishers. Mahwah, NJ, USA.

Dix, A., Finlay, J., Abowd, G., Beale, R. (1998). Prentice Hall Europé.

Gaines, B. R., Shaw, M. L. G., Chen, L. L-Y. (1996). Utility, Usability and Likeability: Dimensions of the Net and Web. In Proceedings of WebNet96. San Francisco, CA, AACE (Association for the Advancement of Computing in Education.

Grose, E, Forsythe, C., Ratner, J. (1998). Using Web and Traditional Style Guides to Design Web Interfaces. In In Forsythe, C., Grose, E., Ratner, J., Human Factors and Web Development. Lawrence Erlbaum Associates, Publishers. Mahwah, NJ, USA.

Laskowski, S., Downey, L., L. (1997) Evaluation in the Trenches: Towards Rapid Evaluation. In Proceedings of ACM Conference on Human Factors in Computing Systems, CHI'97.

Mayhew, D.J. (1998). Introduction in In Forsythe, C., Grose, E., Ratner, J., Human Factors and Web Development. Lawrence Erlbaum Associates, Publishers. Mahwah, NJ, USA.

Nielsen, J. (1993). Usability Engineering. Academic Press.

Nielsen, J. (1999). User Interface Directions for the web. In Communications of the ACM. Vol.42, No.1.

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

Ratner, J. (1998). Easing the Learning Curve for Novice Web Users. In Forsythe, C., Grose, E., Ratner, J., Human Factors and Web Development. Lawrence Erlbaum Associates, Publishers. Mahwah, NJ, USA.

Redmond-Pyle, D., Moore, A. (1995). GUIDE - Graphical User Interface Design and Evaluation - A Practical Process. Prentice Hall Europe.

Schneiderman, B. (1997). Designing information-abundant web sites: issues and recommendations. International Journal of Human-Computer Studies. No. 47, (pp.5-29)

Spool, J., Scanlon, T., Schroeder, W., Snyder, C., DeAngelo, T. (1999) Web Site Usability: A Designer's Guide. Morgan Kauffman Publishers Inc.