

# Join the Joyride!<sup>1</sup>:

## An Identification of Three Important Factors for Evaluation of On-line Entertainment.

Charlotte Wiberg  
Center for digital business  
Department of Informatics  
Umeå University, 901 87 UMEÅ, Sweden  
+46 90 786 6820  
colsson@informatik.umu.se

**Abstract:** This paper identify three different categories of entertainment in the Eurovision Song Contest web site. (1)Contest related content (2) Value-adding content (3). Design- and technology related entertainment. These types arose in an empirical study, a usability evaluation, conducted on twenty test persons. Methods used were usage of general forms, think-aloud protocol with task analysis in combination with a more 'free surf' approach. The level of intervention of the evaluators were also changed between tests. Finally, interviews, both about the web page and the experiment in itself were conducted. The tests were done with both single users as well as pair of users. The main purpose of the tests were to find guidelines in how to evaluate entertainment web sites. The findings show that there are actually guidance to give from out of the three categories of entertainment. Intervention could actually be a good thing sometimes, working together is both good and bad, depending on which of the above mentioned category we are striving to say something about and finally traditional task analysis is not all bad for testing entertainment after all.

### Keywords

WWW, usability, evaluation, experience, entertainment, e-commerce

### Introduction

Entertainment is a big trend on World Wide Web (Jensen, 2000). This paper presents findings from an evaluation of an entertainment- and experiment site. The project is called *Joyride!* and is part of the *Center for digital business*. The project is ongoing, and this paper reports from the first of three phases. Overall, the project will contain 180 user tests. For this first phase, 60 of those have been conducted. The first phase was conducted upon three entertainment web sites. This paper focuses on one of these sites, *Eurovision Song Contest (ESC)*, a web site built to support the TV-event broadcast worldwide in spring 2000, when Sweden hosted this contest. The overall purpose of the project is to develop new methods and criteria for measuring users' reactions to such web sites as entertainment - and experience web sites. The purpose of this paper is to describe the site and the experiment design and give some preliminary findings for how we can categorize entertainment web sites, in order to bring guidance in how usability tests of entertainment sites could be designed in the future. During the evaluation a number of different methods were used, as outlined below. We will try to give some guidance in questions like; What kind of result does the different test approaches give? What combinations of techniques are fruitful in such evaluations? What grade of intervention is appropriate? Should we use single-user tests or let the users work in pairs?

---

<sup>1</sup> The author wants to acknowledge Per Gessle, author of the song 'Join the Joyride', performed by his group Roxette.

The paper contains a broad overview of other work done in usability testing of entertainment and experience systems. More specific definitions of some core concepts related to the problem area is done. Further, an empirical overview is given as well as a description of the design of the evaluation site. Finally, conclusions are drawn from the findings of the tests, and future work in the next two iterations are described and discussed.

## **Related work**

Related work are found in traditional usability engineering (c.f. Nielsen, 1994). Overall, there has been a lot of usability studies made focusing on measuring effectiveness of different types. However, there is a need for new guidelines for designers as well as evaluators of web sites, for building successful entertainment sites. The device "the faster the better" does not work as a measure of success when it comes to such environments (Olsson, 2000). A lot of work has been done on web usability as well, however here a big focus on *information retrieval* is put (c.f. Spool, 1999). Furthermore, the concept of the new economy, by some called "The experience economy" are being explored by for instance Pine II and Gilmore. They give great guidance in exploring the concept of *experience*, and they present fruitful theoretical frameworks in order to bring some focus in search of guidance. However, these authors do not have the intention of further work in evaluation of entertainment sites. They rest on a more general level of abstraction. There is a research field called *Affective Computing*. Rosalind W. Picard wrote a book with the same name (Picard, 1998). However, she is more into agents and person-like interfaces. This research field in general explores the notions of emotions and machines overall. This field has good potential of conducting answers on similar questions. One example of author working with similar questions is Pat Jordan. In his book *Designing Pleasurable Products* he gives guidance in different types of pleasures (Jordan, 2000). Also, he gives guidance in choosing methods for evaluations. However, in that book we get no empirical examples of how the methods work. Also, from our perspective the book is a bit general. It is a great book for argumentation that traditional usability is not enough however. In relation to the above mentioned related work it seems as there is a lack of empirical findings in guidance in evaluation of entertainment web sites, more specifically. It is our intention to fill this gap.

## **User tests**

The user tests were planned in cooperation with the design company. They were interested in differences as well as similarities between the sites. Also, they had interest in the evaluation part. They provided fruitful feedback in the underlying purpose of the designs as well as the focus groups of users. This user groups were used in the finding of test persons. The focus group for *Eurovision Song Contest* was "*Adult person interested in popular music in general and more specifically Eurovision Song Contest*".

## **The test persons**

Inquiries of participation were sent out by e-mail to approximately 60 persons. The test crew sent it to most of the adult they knew in the local area. The issue of ESC interest, the crew left to the test persons to decide. This was not considered when choosing e-mail addresses. Among the group that got the e-mail, two groups could be spotted, i.e. colleagues at the department of informatics at Umeå University and a minor group of informatics undergraduate students. The rest were more mixed in their profile. From these, the answers were handed on a 'first come - first serve' basis. In the end it showed out that all the interested could be tested. The number of tested users for ESC were overall 20. The reward for the tested users was a ticket to the cinema.

## **The test site**

The tests were conducted in a computer lab in connection to the department of informatics. A PC was used as mostly of crowd of test persons had stated in answer form that they were PC-users. The test was conducted on a T3 connection. It could be argued that this type of sites are difficult to download when using a 56,6' modem. However, that was not something we had as a purpose of testing. Nevertheless, it is an important issue. A digital video camera was used to tape the think-aloud part of the test and a mini-disc player was used to record interviews. The video camera was focusing on the user mostly, but also covered some of the screen. The main purpose of the video camera, however, was to get the audio. Also, all tests were conducted with two evaluators. The test crew circled among the activities, but were always present at every time. This so the person not being active could evaluate both the test itself as well as the other evaluator. This really helped in the analysis of the data.

## **Methods used**

First, the users filled in a general form, stating for instance age and interest of ECS. For the latter, a scale from 1 to 7 was used. At the site, another form was used. Here the user answered questions also of very general character.

Examples of questions are *"How often do you use the web"*, *"Do you use a computer at your work"* and so on. These stages all of the users went through. After this, the users were told to imagine a scenario. This with the main purpose to bring the time back to the day after the contest. Also, within the scenario they were told that they came back after a journey and that they wanted to update themselves in order to get a feeling of what happened the evening before. The purpose of this was solely to get them started and to give them a firm 'push' into the web site.

After this the evaluators used different types of test approaches.

- Some of the test persons did the tests together with a friend - others did it alone.
- Elaboration with the level of intervention from the evaluators.
- Some users got a 'free surf' period to start with - others got tasks at once. All of the users did *some* tasks. No user did *only* 'free surf'.
- The evaluators took turn in who was the main evaluator.
- Also, an iterative approach concerning questions in the interview were used and questions were added as well as taken away.

After the usage session, all users were interviewed. The ones working in pairs did pair interviews. Overall the tests took no more than one hour in total.

## The web site - Eurovision Song Contest

The web site evaluated in this paper is, as said before, the support site for the TV-event with the same name. In 2000 the contest was held in Stockholm, Sweden, and the web site was built by Paregos AB<sup>2</sup>. Quoting the description of the web site from the corporate site of Paregos<sup>3</sup>, the purpose of the site was:

*"Swedish Television and Aftonbladet wanted a web site for the Eurovision Song Contest that was not just a pale copy of the television show and they wanted it to present the sponsors in a sensible way. The site were steadily the most visited for the weeks before and after the competition. The visitor can compete in a Song Quiz (with other visitors) and be his/her own DJ by mixing his/her own version of ABBA's Waterloo, and so on."* (<http://www.paregos.com>)

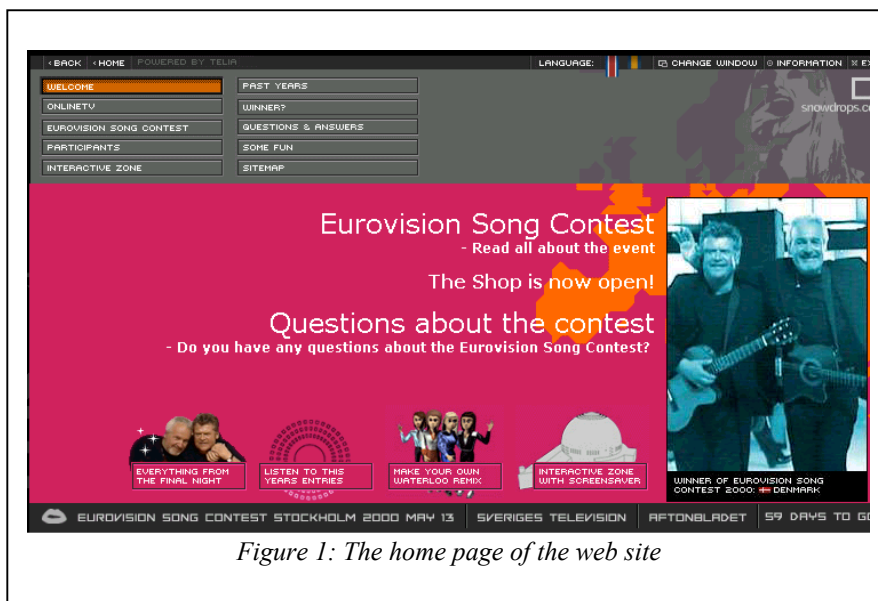


Figure 1: The home page of the web site

The web site uses some plug-ins for the browser as *Macromedia Flash®*, *Beatnick®* and *Shockwave®*. The main menu is found in the upper left. In the very upper corner, a *back* and a *home* button is found. This because the browser goes into 'full screen mode' when the site is starting. At the upper right, changing of language (Swedish or English) could be done as well as information about the design team behind the web site could be found. An *Exit* button is put in the very upper right corner. Further below, in the pink area, there are some high-lighted entries. They end

up at the same pages as in the menu - but they are more specific in their descriptions. In this area there are also

<sup>2</sup> The site is no longer in global use, as it is an event site. However, it can be found at <http://www.paregos.com/eurosong2000/>.

some clickable illustrations, which also are entries to underlying pages. In the very below there are links to external sources. The menu have entries like 'Welcome', 'On-line TV', 'Participate', 'Interactive zone', 'Past years', 'Other fun', 'Site map', 'Ask about ESC', 'News' and more. Under all these entries, lower level entries are found as well. For instance under 'Interactive zone', 'Screen saver', 'mix your own song', 'pop-quiz' could be found. The site will be further explored below.

## Categories of entertainment on the site

When discussing entertainment web sites, the picture often gets a bit blurred. This is because the term 'entertainment' could mean many things. Nevertheless, it is important to grasp this issue to be able to find guidance in how to evaluate such sites. Our empirical findings from the think-aloud evaluations conducted, as well as the interviews of the users, shows that there are actually three categories of entertainment possible to identify on this site. Those categories are more thoroughly described below:



Figure 2: The 'past years' page

1. **Contest related content;** many people have knowledge of these events as a hobby. For instance, there is a club around the contest which you can join and get more knowledge as well as buy material, as a CD with all national winners in Turkey of all times. For some, this content is highly prioritized. Examples of this type of content are; lists of winners of all times, information about this year's contest, video-clips with all the songs,

from parties after the finale. Empirical evidence for identifying this category was for instance, some of the persons

tested, rated very high on the scale of interest of ESC. This type of person rated the contest related content as the best on the site. They were very amused of findings like 'Who won ESC 1981 in Dublin?'. On follow-up questions in relation to this comment, other types of content were used of the evaluators as examples and the person were asked if they, for instance, did not like the 'Waterloo remix'. One person expressed disappointment with this content and said "*it was not as good as the real thing*". **Methodological findings:** Here, traditional task analysis gave good results. Any question seemed interesting and amusing. To answer questions became a game in itself. The users corresponding positively to this type of entertainment on the site saw no difference in how they would handle the contest related content by themselves compared to conducting the test. Interventions as positive feedback when finding right answer seemed important, for highly ESC scoring as those not as interested. This is different from traditional guidelines for task analysis which says that interventions dangers the result *if done during* the task being conducted.

2. **The value-adding content;** The site holds other types of content than such that is strictly related to the content in itself, as downloadable screen savers, ability to send post cards, small games and possibility to remix music. The Waterloo remix (see figure 3) could also be sent to a friend. Also, information of more general kind could be found, as gossip and other light-weight information. This type of entertainment amused the

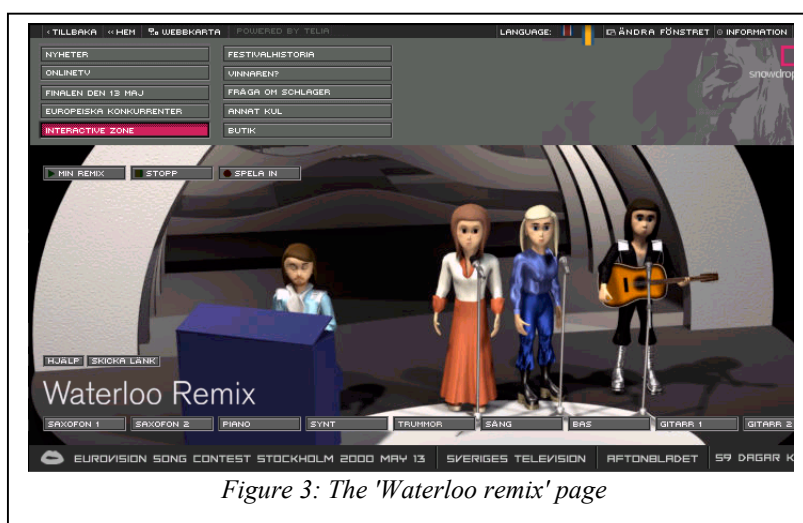


Figure 3: The 'Waterloo remix' page

majority of the tested users. On the question of what was the best of the site, many said 'the Waterloo remix'. They also added that this type of content was the thing that could bring them back to the site or the reason to recommend the site to others. The only ones not responding positively to this content type was the highest and lowest rated on the 'interest for ESC' - the ones in between were positive. Many tested users also said that this type of entertainment was a type they had come across before on the web, and that they sometimes searched for this type of content, in order to test it out - "*play around for a while*". **Methodological findings:** Here, task analysis is less guiding as a method. In order to guide users to the different value-adding features it could be fruitful, but the users given the opportunity to get free surf in the evaluation found these features themselves and 'played around' by themselves. However, note here that some of the users explicitly said that they never explore this type of content on sites and that the tasks therefore not corresponded with how they would react on the site, given that they used it by themselves. The interviews gave feedback for the fact that this was appreciated.

3. **The design - and technical related entertainment:** This type overlaps the second point, as some of the technical features were used as support for more interactive stuff, as games and the remix for instance. However, there are non-interactive animations and likewise on the site as well, which could be put in this last category. Also, all the graphical design is suited to fit in this final category of entertainment.

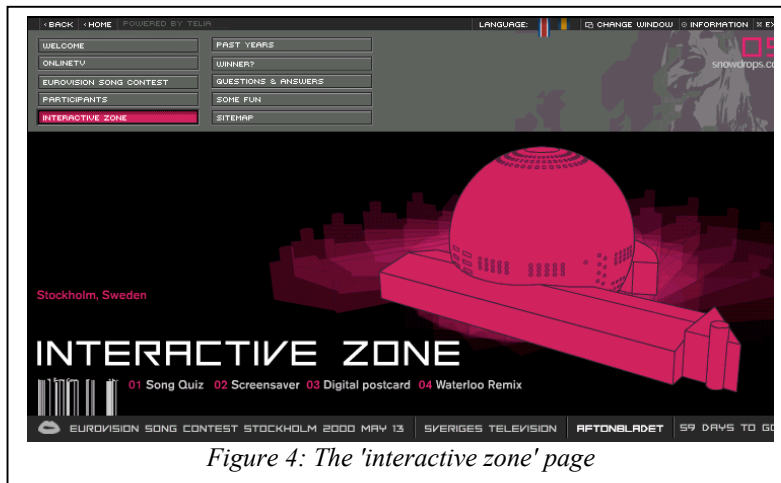


Figure 4: The 'interactive zone' page

This type of entertainment also scored as a factor that amused the users. Spontaneous comments during the 'think-aloud' as well as answers in the interview show that. It seemed natural to talk about "*the design*", "*the form*",

"*the format*" and so on. However, it did not seem as a good idea to separate the static, i.e. the illustrations, the choice of colours and so on, from the moving characters, the animated texts and other type of multimedia features from each other. When asked, users were quite clear that this was the same thing - "*the format seems modern, with the fonts, the colours and the moving stuff*". This type of comment was quite common, and showed upon difficulties of separating the two. **Methodological findings:** Reaction of this type of entertainment was more common from those users tested by themselves. The ones working in pairs reacted on other type of features. Overall, more reflection on details was given by the single testers. Task analysis, in itself does not bring any results at all for this type of entertainment. However, the think-aloud protocol gives guidance. Users self-reflecting about moving features, typography, colours and more. It happened that someone asked "*could I interrupt my activity with the task for a while, I would like to explore this feature*". The tasks disturbed the exploration of this type of entertainment. The interviews, however, gave numerous results in discussions around this category. Intervention from tester is not to recommend, as for the risk of influence.

## Conclusions

In our empirical tests we identified three categories of entertainment on the web site *Eurovision Song Contest*. First, *Contest related content*, like all the top lists of the winners during the years and so on. This was a category amusing the users scoring high on the interest rate of the contest. Second, *value-adding content*, as interactive quizzes, downloadable screen-savers and possibility to mix your own song. This was a type that many of the users mentioned that they usually searches for in this type of sites. Third, and last the *category design - and technical related entertainment*. Examples here are typography, choice of colors and animations of different kinds. Here, no specific user group were spotted. However, it was obvious to the testers that single user tests were more suitable in order to get feedback on the last category. For the first, the *contest related content* task analysis worked fine, which was not true for the second, the value-adding content. Here, users sometimes wanted to interrupt their performance of tasks to explore specific features. Even if it was explicitly allowed, users seemed a bit worried, when interrupting

tasks. Finally, intervention from the testers could be suitable in some of the cases, otherwise it may seem peculiar as a user to explore and being entertained totally alone with someone watching silently with a notepad.

Overall, a great situatedness in tests of this type of web sites seems very clear. The type of test person, elaboration with interview questions and more is important. The users are overall less objectified in testing entertainment web sites compared to traditional usability testing.

### **Future work - progress in the project**

The project *Joyride!* continues with completion of the first iteration as well as the performance of the rest of the 180 tests. This study of *Eurovision Song Contest* gave some guidance but also, as always, enlightened some questions: How about the interventions - *are* they really necessary in some circumstances? What should we do to overlap the gap when task analysis is not suitable any more? Is 'free surf' really the proper solution to that? Are these categories of entertainment general - or will we find more categories further on? Hopefully, future tests will bring us closer to some of the answers of these questions.

### **Acknowledgements**

The author would like to thank the members of Center for digital business at the department of informatics, Umeå University for valuable comments on earlier versions of this paper. Special thanks to Kalle Jegers, Klaas Trojkander and Johan Tufberg, my co-testers in the project, for valuable comments and endless enthusiasm. Finally, I would like to thank Mikael Wiberg for valuable comments on earlier versions on this paper.

### **References**

- Jensen, J.F. (2000) Trends in Interactive content & Services. In proceedings of WebNet2000, San Antonio Texas (pp. 281-286)
- Jordan, P. (2000) Designing Pleasurable Products: An Introduction to the New Human Factors
- Nielsen, J. (1994) Usability Engineering. Morgan Kauffman Publishers Inc.
- Olsson, C. (2000) To Measure or Not to Measure: Why Web Usability Is Different From Traditional Usability. In proceedings of WebNet2000, San Antonio Texas (pp. 425-430)
- Pine II, B. J., Gilmore, J. H. (2000) The Experience Economy: Work is Theatre & Every Business a Stage. Harvard
- Picard, W. R. (1998) Affective computing. The MIT Press. Boston, Massachusetts
- Spool, J., Scanlon, T., Schroeder, W., Snyder, C., DeAngelo, T. (1999) Web Site Usability: A Designer's Guide. Morgan Kauffman Publishers Inc.
- business school Press, Boston, Massachusetts