

Department of Computer and Information Science
Cognitive Science
LIU-IDA/KOGVET-G--11/006--SE

Modifying Heuristic Evaluation for assessing the usability of TV- interaction devices

Eva Bjerke

Supervisor: Magnus Bång

Examiner: Arne Jönsson

Acknowledgements

Throughout the whole project many people have supported and helped me. Initially I want to say I am grateful to Motorola for allowing me access to do this thesis for them. A special thanks are pointed to Otto Carlander at Motorola whom spent many hours on helping me with the project and for his commitment. I also would like to thank Sofia Bremin for being the best partner in crime. Magnus Bång at the University of Linköping also deserves special thanks for providing me with lots of tips and good ideas during the whole project. I would like to thank my beloved Robin Belanger Jensen for his support and for styling a figure. Finally I want to wish special thanks to my dad Kenneth Bjerke and to John Belanger for proofreading.

Abstract

There are a several methods to evaluate usability of systems with graphical user interfaces (GUIs). However, effective methods for evaluating non-GUI interaction devices in the domain of Interactive Television are presently not available. This thesis presents a modified Heuristic Evaluation method for rapid inspection of non-GUI TV-interaction devices such as remote controls. Additionally, to enable the evaluators to more easily think from a user perspective when performing the evaluation, the persona method was also evaluated for use in this domain. The modified Heuristic Evaluation method was evaluated in an actual development project where engineers applied the method on remote control prototypes. The result suggests that the method can be used effectively by engineers and that it identifies usability problems appropriately. The persona approach seemed to provide little support to the engineers in terms of evaluating this type product.

Table of Contents

1. Introduction.....	9
2. Background.....	11
2.1. Interactive Television	11
2.2. Heuristic Evaluation	11
2.3. Persona.....	17
3. Method.....	19
3.1. Research process.....	19
3.2. Modifying the Heuristic Evaluation method	20
3.3. Evaluating the modified Heuristic Evaluation method.....	20
3.4. Our Persona	21
4. Result.....	23
4.1. Modifying the Heuristic Evaluation method	23
4.2. Evaluating the modified Heuristic Evaluation method.....	26
4.3. Our Persona	27
5. Discussion.....	31
6. Conclusion	35
References.....	37
Appendix 1: Questions asked to the stakeholder.....	39
Appendix 2: Template of interview questions.....	41
Appendix 3: Behavior variables	43

1. Introduction

The usability of a product is an important aspect when developing new products. Product developers are different from users in many ways, including their general experience in using a specific product as well as their understanding of the design for the product. Consequently, the developers may think of their product as being perfectly straight-forward and easy to use even though to a user, without the same understanding, it could be completely incomprehensible (Nielsen, 1993). If the product is not designed with the potential users in mind it is likely the product will be rejected. More and more companies are realizing the urgent need for usability evaluations to improve their products' interfaces in order to make the products easy for the users (Nielsen & Mack, 1994). Still, many of the existing usability evaluation methods in use today are too expensive, too difficult and too time consuming to be useful for industry (Nielsen & Mack, 1994). They seem to be developed only to work in theory. Furthermore, the existing methods are almost exclusively developed to evaluate a system interface and none of the studied methods are used for evaluating specific devices. Effective methods for evaluating non-GUI interaction devices in the domain of Interactive Television are missing.

The objective of this thesis is to present a modified Heuristic Evaluation method to evaluate devices for TV-interaction. The method will be tested on a prototype of a TV-interaction device using Personas as a complement to describe the products potential user. This thesis contributes to the following in the area of TV-interaction:

- A modified Heuristic Evaluation method for rapid evaluation of TV-interaction devices
- An evaluation of the modified Heuristic Evaluation method by engineers working in an actual development project on remote controls
- An evaluation of the use of the Personas as a complement to the modified Heuristic Evaluation

The stakeholder is a company that develops products within the area of IPTV. In order to become a part of their future product development cycle, the modified method needs to suit the industry's and the stakeholder's existing development cycle. To achieve this goal the modified Heuristic Evaluation method needs to be:

- Easy to perform for the product developers
- Cost-effective
- Rapid

This thesis begins with a presentation of a short study of TV-interaction. It also proposes Heuristic Evaluation, a method for usability evaluation and Persona, a method to describe the user. In the following chapter the research process will be described and the methods will be explained for how they were chosen. Also, how the Heuristic Evaluation was modified to serve our purpose and how the method was evaluated is described, as well as how the persona was created. Next, the modified Heuristic Evaluation, the result of the evaluation and the persona will be presented. This will be followed by a discussion in which the methods and their results are discussed.

2. Background

This chapter introduces Interactive Television and discusses what services today's remote controls need to handle. Moreover the chapter also presents Heuristic Evaluation, a method for usability evaluation, and Persona, a method for describing a products potential user.

2.1. Interactive Television

Gawlinski (2011) writes that Interactive Television can be described as technology which lets the viewers engage in a dialogue with the TV set. By allowing the viewers to make choices and take actions, their experience goes beyond the passive way of watching TV. Technology development has enabled this capability and advances in digital transmission technologies have made it possible to press a lot more information into a given piece of bandwidth. A common technology used for this is IPTV, a system for transmitting digital TV by means of the Internet protocol, presented to the viewers through a graphic interface (Digital Access, 2006). Examples of services for IPTV are Video On Demand, Pay Per View, Instant Replay and Personal Video Recorder (Digital Access, 2006).

This development has in turn increased the demands on the remote controls. The use of the TV set for Internet surfing further fuels this demand, coupled with a need for text input using the remote control. The increasing services for the remote controls to handle have consequently also increased the need for the remote controls to be easy for the users. A problem when developing remote controls is the broad user group. Most people uses them and they therefore needs to be easy to use both for younger and elder people with different knowledge and interest in technology.

2.2. Heuristic Evaluation

Heuristic Evaluation is a method for finding both severe and less severe usability problems in a user interface design. It is easy to use and since it does not include real users, the method is very time- and cost- effective. On the basis of ten principles, or heuristics, a small set of evaluators inspect the interface and try to find out what is good and what is bad about it. The method and the heuristics were introduced by Jakob Nielsen and Rolf Molich (Nielsen & Molich, 1990; Nielsen & Mack 1994). Even if it is not necessary, it is advantageous if the evaluators know about usability

(Nielsen, 1994b). If no usability experts are available it can be effective to use technical documentation writers as an alternate to evaluate the interface. They generally have more understanding of the system and know when something may be difficult to explain. In such a case the system is usually also difficult to use (Nielsen, 1993).

The method does not provide solutions to fix usability problems found in the interface. Instead the method explains each usability problem with reference to the established heuristics. Once the usability problems are detected they, in most cases, have obvious fixes (Nielsen & Mack, 1994). The Heuristic Evaluation method is well suited for use early in the usability engineering lifecycle since the evaluators do not necessarily need to perform real tasks during the evaluation (Nielsen, 1990).

Nielsen (1994b) describes how experience from many different projects has shown that it is difficult for a single evaluator to do a Heuristic Evaluation since one person can almost never find all the usability problems in a given interface. In the article “Finding usability problems through Heuristic Evaluation” (1992) Nielsen describes a case study where 19 evaluators were used to find 16 usability problems in a voice response system. The study showed that some usability problems were easy to find for any evaluator but some problems were found by only a few evaluators. The result also showed that some of the hardest-to-find problems were found by evaluators who did not find a majority of the usability problems. It is also clear that different evaluators find different usability problems. Based on this, Nielsen (1994b) recommends a normal use of three to five evaluators, since increasing in significant information are not gained by using more evaluators.

How to perform a Heuristic Evaluation

In a Heuristic Evaluation each individual evaluator inspects the interface alone. It is very important that the evaluators do not communicate or aggregate their findings before they all have completed their evaluations. This is to ensure independent and unbiased evaluations from each evaluator (Nielsen & Mack, 1994).

During the evaluation, the evaluators either can write down their findings themselves or tell an observer who writes them down during the session. When using an observer, individual evaluation sessions will take more time but this also gives the evaluators

more time to focus on the system (Nielsen 1994b). The result of the evaluation will also be available faster because the observer only needs to aggregate his or her own notes; not a set of reports written by the different evaluators. The observer can help the evaluator with problems such as unstable prototypes or to explain certain aspects of the interface during the session. By answering the evaluators' questions about the overall functions of the prototype, the observer enables them to focus on the usability of the user interface rather than wasting time struggling with, for example, navigating the interface (Nielsen 1994b).

A typical Heuristic Evaluation session normally takes between one and two hours for each individual evaluator. If the system is very complex, longer evaluation sessions may be needed. In that case it is usually better to split the evaluation into several smaller sessions; one for each part of the interface (Nielsen & Mack, 1994). During a session the evaluator goes through the interface inspecting different parts and compares them to a list of usability heuristics. The evaluators can then decide how many times they want to go through the interface. A general recommendation is at least twice. The first run-through is for the evaluators to get to know the flow of the interaction and the system. The second session allows the evaluator to focus on specific elements of the interface, in order to find out which elements they should use in the whole system (Nielsen & Mack, 1994). Supplying the evaluators with scenarios consisting of realistic tasks makes it easy for them to learn the system; which is also helpful during the evaluation (Nielsen, 1990). The scenario should be constructed based on the needs of the actual users and listing the actual steps the users would take to perform those tasks and thereby to be as representative as possible (Nielsen, 1993).

Nielsen and Molich (1990) developed the first, and still most common, heuristics that they believe should be adapted by all user interface designers. Nielsen later (1994a), based on a factor analysis of 249 usability problems, refined their heuristics and came up with these ten heuristics:

- **Visibility of system status**

The system should always keep users informed about what is going on through appropriate feedback within reasonable time.
- **Match between system and the real world**

The system should speak the users language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.
- **User control and freedom**

Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.
- **Consistency and standards**

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.
- **Error prevention**

Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.
- **Recognition rather than recall**

Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

- **Flexibility and efficiency of use**
Accelerators -- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.
- **Aesthetic and minimalist design**
Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.
- **Help users recognize, diagnose, and recover from errors**
Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.
- **Help and documentation**
Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

Nielsen's heuristics are broad and apply to any type of user interface, both character-based and graphics based. Nielsen (1994b) points out that the evaluators, if necessary, can change or add to the heuristics to suit their specific interface.

The result of a Heuristic Evaluation method is a list of usability problems with the interface and identification of the specific heuristic each problem infringes with. The evaluators need to motivate and explain what they do not like. They should be as specific as possible and each usability problem should be listed separately. If there are, for example, four problems with the same part of the interface, all four should be described individually with reference to the heuristics. This is to make it as easy as possible for the designers to fix the problems (Nielsen & Mack, 1994).

To determine which of the found usability problems are the most severe, the evaluators can grade them. This helps to identify which problems need the most

resources to have them fixed and assists in deciding upon which problems to give priority to and vice versa.

The severity of usability problems is a combination of three factors (Nielsen & Mack, 1994):

- The frequency with which a problem occurs. - Is it common or rare?
- The impact of the problem if it occurs. - Will it be easy or difficult for the users to overcome?
- The persistence of the problem. - Is it a one-time problem that users can overcome once they know about it or will users repeatedly be bothered by the problem?

As part of the evaluation process, each evaluator should grade all usability problems, including those found by other evaluators, not just the problems identified by themselves. This can be easily be done by a questionnaire with the complete set of problems and a five-point rating scale for each problem (Nielsen & Mack, 1994):

- 0** I don't agree that this is a problem at all
- 1** Cosmetic problem only- need not be fixed unless extra time is available on project
- 2** Minor usability problem- fixing this should be given low priority
- 3** Major usability problem- important to fix, should be given high priority
- 4** Usability catastrophe- imperative to fix this before product can be released

To broaden the Heuristic Evaluation method into providing some design advice, a debriefing session can be held after the last evaluation session and the severity questionnaire. The participants in the debriefing should be the evaluators, the observer and some people from the design team. The debriefing is a brainstorming session with focus on discussions of possible redesign and solutions for the usability problems (Nielsen & Mack, 1994).

2.3. Persona

In order to make the product successful each member of the development team must understand the fundamental characteristics and needs of the customers and users (Goodwin, 2009). There are a number of methods to help the developers gain insight of the users needs. Persona is one of them.

A persona engages the social and emotional aspects of our brain to help the team members to visualize the best product behavior and to see the why the recommended design is good. A persona is a fictive person that represents the potential users and describes their various objectives and observed behavior patterns. The method is used to encourage engineers to think from a less-skilled user's perspective and was first introduced in 1995 by Alan Cooper (Goodwin, 2009). Personas uses storytelling to explain the most critical behavior information in a way that designers and stakeholders can understand. By gathering information from and about the user through interviews and by analyzing the information a story that includes name, a photo, a set of goals, environment, skills, frustrations, attitudes, typical tasks, and any other factors that seem critical in understanding the users behavior pattern is made. Each persona has its own goals and needs. It is often necessary to create at least two personas since you almost always find two or more diverse types of opinions or behavior amongst potential users (Goodwin, 2009).

3. Method

In this chapter the research process is being described. Moreover, how the Heuristic Evaluation were being modified and evaluated is being presented, as well as how the Persona was created.

3.1. Research process

This chapter describes the research process of the project. How and why the methods were being chosen and modified to evaluate TV-interaction devices. To clarify what the stakeholder expected of the project the work started with an interview with the stakeholder. For all questions see Appendix 1. We learned that the methods for evaluation should be applicable on remote controls and keyboards with only a limited aspect of GUI since our stakeholder does not design GUI. Focus should not be held on the esthetic of the remote control but on the demands the users have on it and which demands the remote control puts on the users. We also found out that the potential users are an “adder”, since the product is not a standard product. The users choose to buy it because they are interested in using the IPTV services. The users are assumed to be people in the age of 18-45. 75% are likely men. One other important aspect we found was the low budget our stakeholder works with on performing evaluations in the future. The company has limited time to perform the evaluation and analyze it and one person must be able to do it on his own. This was seen as the most important aspect in choosing the right methods.

Further, a brief study of some usability evaluation methods was being made. According to the stakeholders' requirement a table with every advantage and disadvantage of each method was set up. The most advantageous methods, Usability Testing, Cognitive Walkthrough and Heuristic Evaluation, were suggested to the stakeholder. Persona was suggested as a method to describe the potential users to help the evaluators to think from a user perspective when performing the evaluation. Together with the stakeholder a discussion were being held to establish this methods are the most suitable to use. Since Usability Testing claims real users, which can be hard to find, it was excluded in consensus with the stakeholder. Followed the persona was being created and the chosen evaluation methods were being modified to suit our specific type of product.

This thesis only describes Heuristic Evaluation and Persona. For the Cognitive Walkthrough method see Sofia Bremin's "A modified Cognitive Walkthrough method: for evaluating TV-interaction devices" (In press).

The figure below shows the elements of the research process in order step by step.



Figure 1. The research process

3.2. Modifying the Heuristic Evaluation method

Since the goal of Heuristic Evaluation is to find usability problems mainly in a graphical user interface design (Nielsen & Mack, 1994), the method had to be modified in order to be applicable to the evaluation of a remote control with no graphical user interface.

The modification was done by reducing the ten heuristics to eight. The remaining eight heuristics were then rewritten to become simplify, accessible and oriented to remote controls. Further, instead of each evaluator grading the found problems by themselves, the method was changed so that all evaluators discuss and grade all problems together. Finally one instruction for the coordinator and one for the evaluators was written. The instruction was composed to be easy to read and understand for an inexperienced evaluator.

3.3. Evaluating the modified Heuristic Evaluation method

To assess if the modified Heuristic Evaluation effectively fulfilled its purpose, an evaluation of a remote control prototype was held together with the stakeholders.

Three evaluators participated, one of the stakeholders and two students. The stakeholder had almost no knowledge of the method and the students were those involved in this study and responsible for the modification of the evaluation method. One by one the evaluators began by reading the instruction sheet about how to perform the modified Heuristic Evaluation. They then inspected the remote and compared it to each heuristic.

Two of the evaluators wrote down their found problems themselves, explained and motivated which specific heuristic the problem infringed with. Responses were recorded on the paper listing the identified heuristics. The two students did their evaluation separately from each other since the evaluators were not allowed to discuss or aggregate their findings together before they performed their own individual evaluation. The third evaluator provided his findings to one of the students who at that point acted as an observer and wrote them down.

When all three had finished evaluating the remote control, one of the students took the role as coordinator and compiled all problems found by all evaluators. Together all three evaluators discussed each found problem and graded it using the scale on the information sheet to get the severity of each problem.

They also discussed possible redesigns and improvements on the remote control. After the session the stakeholder redesigned the prototype to fix the most severe problems. A couple of days later all three evaluators sat down again and discussed the new prototype and compared it to the most severe problems found earlier. This was to see if the problems were fixed and if the remote control was improved.

3.4. Our Persona

To create the persona, nine persons were interviewed, six men and three women. All of them had different educational background and were between the ages of 23-49. They all had different levels of knowledge and interest in TV-interaction and new technology. Seven of them were employees at our stakeholder with different occupations. One was a student and one was self-employed.

Semi-structured interview techniques were used with a base template of questions (Appendix 2). The template of questions consisted of 16 predetermined questions with the aim to get to know the participants' interest in technology as well as their behavior and attitudes to TV- interaction and TV- devices. Each interview took approximately 10-20 minutes. Seven of the interviews were held by two persons; one person asked the questions and the other took notes. Those interviews were held at a meeting room at our stakeholder.

The remaining two interviews were held by a single individual who both asked the questions and took notes. These interviews took place at the participants' homes.

All participants were also told that the interviews were recorded and that their answers would be kept confidential. One participant elected to be recorded.

The answers from the interviewed persons were presented with numbers and were compiled and categorized into behavioral and demographic variables (Goodwin, 2009). Patterns between the answers were found and marked up with different colors. By studying the differences between the participants' behavior and attitudes enough data was available to create one persona.

4. Result

This chapter describes the modified Heuristic Evaluation as well as the result of the evaluation of the method. Also the Persona is presented.

4.1. Modifying the Heuristic Evaluation method

The heuristics in the original Heuristic Evaluation are formed to evaluate a system with a graphical user interface. To suit our purpose, to evaluate a remote control without a GUI, these two system-concentrated heuristics were omitted:

- **Visibility of system status**
The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.
- **Help users recognize, diagnose, and recover from errors**
Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

The remaining eight heuristics were rewritten and modified hence the original heuristics were too general and hard to understand. We modified them to be specific for evaluation of a device and also to be accessible. These are shown below:

- **Match between device and the real world**
The device should speak the users language, with icons, labels and concepts familiar to the user, rather than device-oriented terms. Follow real-world conventions by making information appear in a natural and logical order.
- **User control and freedom**
Users often choose functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.
- **Consistency and standards**
Users should not have to wonder whether different labels, icons, or actions mean the same thing.

- **Error prevention**

Is the device designed carefully to prevent problems from occurring? Are all buttons placed in such way users cannot press them by mistake? Is each buttons' intended function clear? Is each buttons' label placed intuitively?
- **Recognition rather than recall**

Minimize the user's memory load by labeling buttons and making actions and options visible.
- **Flexibility and efficiency of use**

Allow users to tailor frequent actions. Shortcuts may often speed up the interaction for the experienced user and make the device suit both inexperienced and experienced users.
- **Aesthetic and minimalist design**

The device should not contain information that is irrelevant or rarely needed. Every extra unit of information competes with the relevant units of information and diminishes their relative visibility.
- **Help and documentation**

Even though it is better if the device can be used without manual, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

Below you will find the instruction sheet for the coordinator and the evaluators on how to perform the Heuristic Evaluation. The coordinator is the one who is in charge of the evaluation.

How to perform Heuristic Evaluation

For the coordinator before the evaluation:

Use 3-5 evaluators. Each individual evaluator shall go through the device alone. It is very important that the evaluators do not speak to each other or aggregate their findings before they all have completed their evaluations.

During the evaluation the evaluators either write down their findings themselves or they give their comments to an observer who writes them down during the session. The observer can help the evaluator with any problems such as unstable prototypes or to explain certain aspects of the interface during the session.

For the coordinator after the evaluation:

Compile the problems found by all evaluators and gather all evaluators. Discuss each problem in order to assess the severity of the problem. The severity of a usability problem is a combination of these factors:

- The frequency with which a problem occurs. Is it common or rare?
- The impact of the problem if it occurs. Will it be easy or difficult for the users to overcome?
- The persistence of the problem. Is it a one-time problem that users can overcome once they know about it or will users repeatedly be bothered by the problem?

Grade each problem, using this scale:

- 0** This is not a problem at all
- 1** Cosmetic problem only- need not be fixed unless extra time is available on project
- 2** Minor usability problem- fixing this should be given low priority
- 3** Major usability problem- important to fix, should be given high priority
- 4** Usability catastrophe- imperative to fix this before product can be released

For the evaluators:

Inspect the device and compare it to each heuristic on the list of usability heuristics. Write down or tell the observer which problem you found and which heuristic each problem infringes with. Motivate and explain why you do not like it. Be as specific as possible. List each problem separately. If there for example are four problems with the same part of the device, all four should be described with reference to the heuristics.

4.2. Evaluating the modified Heuristic Evaluation method

The problems found during the evaluation of the modified Heuristic Evaluation and which heuristics they infringe with is shown below. The number indicates the severity of the problem using the scale shown in the method chapter for modifying Heuristic Evaluation.

Note: Since the remote control prototype used in the evaluation is under development and considered company proprietary, not all of the found problems are presented in this report, as some of them were considered to be too product specific.

- **Match between device and the real world**

Text instead of symbols on some buttons	1
Some buttons are not in their regular place according to users from the Nordic region	0
The color buttons do not follow the standard	3

- **User control and freedom**

There is the “TV”-button but it takes a giant step back in the menu	0
There is “Backspace” but it’s marked with text not arrow	0
There is a “Back”-button but it’s placed slightly invisible	3

- **Consistency and standards**

“Shift” and “Control” are placed twice	0
The “Shift” buttons are not in the same size	0
There are several “Quotation mark” buttons	0
The “Enter” and “OK”- buttons have the same function?	1

- **Error prevention**
 - The small buttons may be placed too tight 3
 - Harder to push the buttons with the left hand than the right 2
 - “Alt Gr” + the color buttons may be tricky to understand 3
 - “Menu” is placed invisibly 3

- **Recognition rather than recall**
 - All buttons are easy to find after they have been used once 0

- **Flexibility and efficiency of use**
 - The “TV”-button is a shortcut
 - Shortcuts can be used but there is no opportunity for programming 0

- **Aesthetic and minimalist design**
 - Some symbols may be used very rarely 0
 - The “Text”-button has the best place, will it be used that often? 1

4.3. Our Persona

Examples of variables found by analyzing the data from the interview were personal information such as age, education and family. Variables such as interest in buying new technology, how much time spent on watching TV were also identified. We also found out if services like Video On Demand, Pay per view and Personal Video Recorder were used and what experience the interviewed subjects have in using a keyboard for surfing on TV. Variables about demands on remote controls and negative attitudes about the remotes were also seen. For all behavior variables see Appendix 3.

Examples of the participants negative experience of remote controls were among others, bad layout, substandard functionality, bad ergonomics and that the devices are not designed in an intuitive manner. These findings corresponded to their demands of functionality for remote controls.

According to the behavior variables this persona was created:

Stefan Lindgren

Stefan is 38 years old and works as a project leader at Ericsson in Linköping. He was born and raised in Karlstad, but moved to Linköping to study engineering in computer science. He lives in a villa together with his wife and two kids, who are four and six years old.



Stefan is quite interested in new technology and mainly in technology used in everyday life. He thinks that it is important to be aware of the latest technology, but does not feel the need to buy the latest products. Stefan and his family have two TV-sets at home. Sometimes the family watches TV together and sometimes separately. It is mostly Stefan's wife who decides what program they will watch. Stefan himself watches TV about 1-2 hours a day. The family has the large package from ComHem and they use the digital TV-box mainly for renting movies and recording programs and movies.

Stefan has no experience in surfing the net using the TV, or using a keyboard in interaction with the TV. Although he can imagine himself using a keyboard in the future.

Stefan's opinions about his current remote controls are that they have too small displays. Some buttons are also either too big or too small. One thing that he finds particularly annoying is when the remote controls react poorly when he presses the buttons. During the weekends when Stefan and his wife typically watch movies, he finds it difficult to get to the right buttons on the remote controls if the room is dark.

Stefan finds it frustrating to have several different remote controls. If he buys a new one it is important that it supports all of the functions both on the TV-set and on the digital box.

Today the family uses the remote controls followed with the purchase of the products.

Stefan's goals with remote controls:

Design: Stefan wants the remote control to feel comfortable in his hand and the buttons to be placed intuitively.

Functional: The remote control has to react quickly when Stefan pushes a button and support all the functions in his TV and boxes.

Solid: It is important to Stefan that the remote control is not fragile since his small children often drop them on the floor.

Intuitive: Stefan wants the remote control to be so easy to use from the start so that he doesn't need the manual.

5. Discussion

Heuristic Evaluation was an obvious choice when choosing an evaluation method since it approached the stakeholder's demands on being fast and cost effective. Beyond this, the method can be performed by the developers themselves and it can easily be done several times during the entire design and development process. Still, it was a challenge to modify the method into a model where the GUI is not the evaluation focus as it is to with most of the existing methods today.

In fact, the purpose of the modification was not only to make the method suitable evaluating remote controls. We also wanted it to be easy to perform for a stakeholder who does not have any in depth knowledge about the method. One of the difficulties with the original method is the simple fact that the Heuristics can be hard to understand so we focused on trying to make them comprehensible.

The outcome was a less system-oriented method with eight simplified heuristics and an instruction sheet. Thanks to this change, it is credible that evaluators with no usability background would still understand and be able to perform the method. The stakeholder for this study thinks that the method feels natural to perform and that it is applicable within their industry. This implies that we managed to simplify the method.

The result of the evaluation of the modified method shows that it works effectively since several usability problems were, in fact, identified on the prototype. The stakeholder's opinion is that through the use of this method the usability problems can be found at an early stage of the development process. One of the other major benefits is that the method forces the in house evaluators to sit down and really think about their product and evaluate it in a structured manner.

The provision of specific instructions to follow simplifies the process for the evaluators, and prevents them from focusing on single aspects of the device and thereby misses other issues and aspects. A potential weakness with the method is the fact that the evaluators go through the device as a whole rather than its parts or sub functions. Hence looking at the whole it is possible to miss specific parts or details of the device.

The purpose of the persona was to better understand the user and to make it easy for the evaluators to think from the user's perspective instead of their own during the Heuristic Evaluation. This aspect is important since developers tend to think that the users have the same knowledge as themselves, which often is wrong and can be devastating for the usability of a product.

A problem with the persona in this project was the broad group of potential users. It is easier and more useful to make a persona when the potential user group is smaller since the needs and demands on the product are probably more specific. The people that participated in the interview for the persona were too homogenous since all nine were at least further educated. It is preferable to have a better mixture of people when you perform these types of interviews.

All evaluators were acquainted with the persona when performing the evaluation but none had studied the persona in at least the week prior to the evaluation. Perhaps it would have been more useful if the persona had been read just before the evaluation of the remote control. Even if we thought we knew the persona and that we thought from its perspective it would have been easier to do this if it was fresh in our memory.

The Heuristic Evaluation method was modified so all evaluators gathered after we all had performed our evaluations. We discussed all problems together and graded the responses to assess the severity of each problem. We also came up with proposals for redesigns. This activity was very valuable. When the evaluators individually grade the problems they probably easily can do it a bit carelessly and without great consideration. In a group environment, each problem was discussed in detail and each individual's feedback was carefully weighed. This forced us to really think of each problem and the grades of severity were really considered. At the same time, redesigns were being discussed in a natural way.

After using the persona the stakeholder did not find it very useful. Instead he thought it would be better to use interviews or surveys for understanding the users. In our opinion it is not possible to get the same understanding of the users with those methods and they can therefore not replace a persona. However, for this type of product, when the group of potential users is broad, e.g. remote controls, a persona is not optimal. Interviews and surveys can be a better way to obtain the users feelings and attitudes about this type of product. With these methods you can get the users

common needs, demands and problems with the products, which can be helpful when designing for the users. Alternative, to cover to whole group of potential users you could make a few personas instead of just one.

One reason why the stakeholder did not find the persona useful was that the persona reminded him too much of himself. If the persona had been an old lady, it might have been found to be a more useful tool. However, since the potential user for the product of our evaluation is most likely a man between the ages of 18-45 with an interest for technology it somehow shows that the created persona “Stefan” feels realistic which is a positive thing. Since probably most people in the development team at our stakeholder, belong to this group, and therefore are potential users, they probably do not have problems designing for the users for this specific product as they might have when designing for other types of users. A persona can therefore be more helpful when they design for a group that is farther away from themselves.

Seven of the persons in our persona interviews are employees at our stakeholder and this may have influenced their answers. But since they all have different occupations and most of them do not work directly with TV-products, they have no more insight about the products than persons outside of the company. Therefore, it probably did not matter for the result.

Since we do not have full insight in the industry it is possible that we would have made a different choice of evaluation methods if we knew more about Interactive TV. A deeper observation of the stakeholder’s product cycle would therefore be preferable. A possible approach for choosing a more appropriate method could be by observing which methods for evaluation are used by other companies to see what works best in practice. However this can be hard to perform since companies often keep their product cycle and methods classified.

Although the method “Usability Testing” did not fulfill the stakeholders criteria on being time- and cost effective we suggested it because it is often referred to as the best method for usability evaluation and that it is in some way irreplaceable (Nielsen, 1993). We discussed the method with the stakeholder but since it includes real users which are hard to find, as is as well expensive and time-consuming, we decided not to use it. However by including real users direct information about the product is gained, such as if it works the way it is supposed to and if it is easy to use. With Usability

Testing, problems that are not found with other methods that do not include real users are more readily found. Therefore, Usability Testing is overall the preferable method to use.

6. Conclusion

This thesis presented a modified Heuristic Evaluation method to inspect the usability of remote controls for Interactive Television. Moreover, to improve the evaluators' ability to think from a user's perspective when performing different evaluations, a persona-oriented approach was also developed. The result of the modified Heuristic Evaluation indicated that the method is easy to understand and perform for engineers. By providing the evaluators with clear instructions to follow in a structured manner, the method provides an opportunity for developers to find usability problems on their prototypes early in development. We believe that the modified Heuristic Evaluation method will be a good addition to engineering practices in the area of Interactive TV. The result of the persona evaluation showed that the method is not very useful when the user group is as broad as is the case for remote controls. Future improvements of the modified Heuristic Evaluation method could be to include user tasks scenarios. This addition would improve the evaluators' ability to fully assess more functions and aspects of the device.

References

Bremin, S. (In press). *A modified Cognitive Walkthrough method: for evaluating TV-interaction devices*. Linköping: Department of computer and information science, University of Linköping.

:: *DigitalAccess AB :: Triple play solutions*. Retrieved 4/26/2011, 2006, from www.digitalaccess.se/document/Vad_ar_IPTV.pdf

Gawlinski, M. *Interactive television production - a definition of interactive television*. Retrieved 4/26/2011, 2011, from [http://www.interactivetelevisionproduction.com/What is interactive television.html](http://www.interactivetelevisionproduction.com/What_is_interactive_television.html)

Goodwin, K. (2009). *Designing for the digital age : How to create human-centered products and services*. Indianapolis, IN: Wiley Pub.

Mack, R. L., & Nielsen, J. (red.) (1994). *Usability inspection methods*. New York: Wiley.

Nielsen, J. (1990). Paper versus computer implementations as mockup scenarios for Heuristic Evaluation. *Proceedings of the IFIP Tc13 Third International Conference on Human-Computer Interaction*, 315-320.

Nielsen, J. (1992). Finding usability problems through Heuristic Evaluation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 373-380.

Nielsen, J. (1993). *Usability engineering* ([New ed.] ed.). Boston: AP Professional.

Nielsen, J. (1994a). Enhancing the explanatory power of usability heuristics. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Interdependence*, 152-158.

Nielsen, J. (1994b). Guerrilla HCI: Using discount usability engineering to penetrate the intimidation barrier. *Cost-Justifying Usability*, 245-272.

Nielsen, J., & Molich, R. (1990). Heuristic Evaluation of user interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Empowering People*, 249-256.

Appendix 1: Questions asked to the stakeholder

1. What products should the evaluation methods be applicable on? Solely remote controls and keyboards?
2. Where should the focus be of the aspects mentioned below? Are there any other aspects that are more important?
 - Ergonomics, the product's performance, interaction between product and the user? (Could include all)
 - The users feeling for the products appearance? (The products esthetics)
 - Which mental/physical demands the product puts on the user, and vice versa?
 - The competition with other similar products?
3. What is the budget for executing the evaluations in the future? Is this something we need to consider while designing our study?
4. What is the target group that you would like to focus on during the evaluations?
5. Do you have any available users for the evaluations or should we find these ourselves?
6. Is it possible to provide the participants with any compensation?

Appendix 2: Template of interview questions

We are studying the bachelor program of Cognitive Science at the University of Linköping. This spring we are working on our bachelor thesis here at Motorola.

This interview will take about 10-20 minutes.

Your participation is completely volunteer, anonymous and your answers will only be used in this project. You do not have to answer all questions and you can choose to break at any time.

Is it OK with you if we record the interview? The recordings will only be used for supporting our work with analyzing the interview.

If you have any questions about the project feel free to ask after the interview!

1. How old are you?
2. What education do you have?
3. What is your occupation?
4. How many people are there in your household?
5. Grade from a scale of 1-7, how interested are you of buying the latest technology?
6. What sort of technology are you most interested in?
7. Who in your household are in charge of what you are watching on the TV?
8. How much time do you spend on watching TV per day? How do you use your TV?
9. Do you use any extra services (rent movies, internet surf, record programs etc.)?
 - If yes: How do you use these extra services?
 - How often do you use them?
 - What support do you get from the remote control using these services?
 - If no: Why?
 - Would you be interested in using any extra services in the future?
10. Is there anything negative about remote controls?

11. What demands do you have on remote controls?
12. What is the best with the remote control you use today? What do you use it for?
13. Have you replaced the standard remote control for your TV-set against another one? If yes: Why? How do they separate?
If no: Would you like to replace it? What for?
14. Tell what functions you would like to have on a device for surfing, chatting, changing channel, recording and any other TV-interactive services?
15. Do you have any experience in using a keyboard for internet surfing on the TV? If yes: What experience?
16. Can you see yourself buying a wireless keyboard in the future for interaction with the TV? If yes: Tell what you would use it for? Please, motivate your answer.

Thank you for participating!

Appendix 3: Behavior variables

Age

23 (9), 28 (1), 31 (5), 32 (8), 38 (3), 47 (4), 49 (7), 49 (2), 50 (6) (Medium-/mean value: 38 år)

Education

High educated Further educated

3, 2, 4, 7, 6

1, 5, 8, 9

Household

Cohabit Family

1, 5, 8, 9, 6

2, 3, 4, 7

Interest in new technology

High Medium Low

2, 3, 8, 9

1, 4, 6, 7

5

Willingness in buying new technology 1-7

1 2 3 4 5 6 7

2, 7 4, 6, 9 1, 5 3 8

Time spent on watching TV

1-2 h 2-4 h 5-6 h 6 < h

2, 3, 4, 7, 6, 1, 8, 9 5

Have got extra services

Yes No

3, 1, 4, 6

7, 5, 2, 8, 9

Wish to have extra services (Of the ones who do not have extra services) and reasons why they do not have it today.

Yes:

2 (it is too expensive), 7 (it is too expensive) 9 (the supply are too lousy)

No:

8 (it is better to use the computer)

Used the TV for internet surfing with a keyboard

Yes

No

6, 5, 3, 2, 8, 9

7(would like to), 4 (have it but not used it yet), 1(would like to)

Negative things about remote controls

Nothing (1)

You need so many (2)

Ergonomic (5, 3, 8)

Layout (6, 7, 3, 9)

Fragile (4, 8)

Not intuitive (4, 6)

Needless buttons (5)

The buttons are hard to find in the dark (6, 8)

Outwears (6)

Slow reaction (7)

Demands on remote controls

Useful/ Layout (7, 3, 4, 1, 5, 8, 9)

Ergonomic (1, 2, 3, 5)

Functional (6, 2, 4, 1, 5, 8, 9)

Shows in the dark (2)

Sustainable (2)

Reacts well (7, 8, 9)

Has the standard remote controls for the TV-set been replaced

Yes

No

2, 3, 6

7, 4, 1, 5, 8, 9

Desired functions of a remote control for interactive services

Multifunctional (1, 2)

Touch (2, 6)

Instant feedback (2, 4)

Clearness (4)

Keyboard (7)

Keyboard + touch (3, 8, 9)

Would buy a keyboard for interact with the TV

Yes

No

7, 2, 3, 1, 5, 9

No:

6 (I would rather buy a tablet), 8 (I would rather use my mobile phone)