OULUN YLIOPISTO
UNIVERSITY of OULU

Challenges and advantages of using usability and user experience heuristics – Case Facebook

University of Oulu
Faculty of Information Technology
and Electrical Engineering
Department of Information
Processing Science
Master's Thesis
Laitinen Olli-Pekka
17.03.2014

Abstract

User experience became a buzzword in the mid 90's after it was conceived. It quickly took over the previous position of usability as the leading term to cover the interactions between man and machine. Heuristics, originally created to assess usability in systems were adapted to try and evaluate the experience which is drawn from the interaction with the systems. However, usability and user experience differ from one another with many aspects. While the heuristic evaluation is procedurally similar when evaluating usability or user experience, the fundamental differences between the two challenge the applicability of the method itself.

The aim of this thesis is to take a critical look at usability and user experience evaluation using general usability heuristics created by Nielsen and Molich in 1990 and user experience heuristics created by Arhippainen in 2009. This thesis will explore the challenges and advantages that heuristic evaluation of systems usability and user experience encompasses. This is done by reviewing relevant literature and by conducting two case studies for evaluating Facebook's usability and user experience.

This thesis will eventually present that the original advantages of heuristic evaluation of systems usability are mostly applicable to the UX heuristics with minor differences. However some additional challenges concerning the subjective and temporal nature of user experience are discovered when applying the UX heuristics. This thesis will also suggest that using heuristic evaluation to evaluate systems user experience is recommended not alone, if possible, but in conjunction with other methods.

Keywords: heuristic evaluation, user experience, expert evaluation, usability, Facebook

Foreword

This thesis has been in writing for months and I am pleased that it is finally in this stage; up for presenting. The field of research has been astoundingly interesting and I am certain that without the misfortunes during the writing process, it would have been completed sooner. I am however proud of the outcome and I believe that the work I have completed is contributing to the body of literature by providing at least somewhat clearer and unified image on the topic. I want to thank Leena Arhippainen Ph.D. and Anna-Liisa Syrjänen Ph.D. whom both have abided by me and encouraged me during the writing of this thesis.

 in Oulu 17.3.2014

Table of Contents

1. Introduction

Information systems have changed - from traditional view of information systems as utilities designed for a clear purpose into omnipresent services and systems encompassing both work-related utilitarian aspects as well as fun related hedonistic aspects (Thong, Hong & Tam, 2006). The customers have become more powerful than ever before and now get to decide what systems they will use and discard those that do not fulfill their expectations for appropriate and effective use. Ever more so, the failure to fulfill these expectations often contributes to corporate failures. (Lyytinen & Hirschheim, 1987.)

As the world has rapidly moved towards service oriented business ecosystem, and the capabilities of information creation and sharing are ever the highest, even one disapproving voice can make the difference between acceptance and rejection (Oinas-Kukkonen & Oinas-Kukkonen, 2013). Due to this reason, users' approval of products has become a key metric and a prerequisite for success. To assess this approval prior to a product launch, many methods have been created. Different forms for evaluating products usability and thus its acceptance, generic technology acceptance models and their extensions. Despite all the methods available, acceptance of new information systems has been too negligible for too long (Venkatesh & Davis, 2000).

There are services available that work like a charm, they do what they are expected to do, and they do it well at that. But not all of those services make it in the tough ecosystem that is the modern world. How come some services make it even though there might be some minor glitches and drawbacks in their design and those that are built better become forgotten? How did Facebook become so popular even though it failed in a usability research? (Hart, Ridley, Taher, Sas & Dix, 2008) What is the reason that Angry Birds have managed to become a worldwide phenomenon even though it is hardly the first game of its kind? In the midst of this dilemma is rooted the definition of user experience (UX).

Great user experience is something that everyone is aiming at (Hole & Williams, 2007; Väänänen-Vainio-Mattila & Wäljäs, 2009), this is done even unconsciously at times. This makes the research into the topic highly needed and current. Understanding and thus creating a good user experience is considered a critical task. However there are difficulties in understanding and more importantly quantifying experience (Forlizzi & Battarbee, 2004).

Even though user experience research suffers from multitude of different definitions (Appendix A) and opinions, the subject is far from mature, researchers agree on one thing – the subject of research is far too important to dismiss. (Law, Vermeeren, Hassenzahl, & Blythe, 2007). Therefore we can assume that methods to assess and evaluate user experience are equally as important and there is a dire need for research on such methodologies. For this specific reason, Arhippainen (2009) created the general user experience heuristics to apply the method created by Nielsen & Molich (1990) to evaluate systems user experience. However user experience is a much wider relationship between the system and a user (Law, Roto, Hassenzahl, Vermeeren & Kort, 2009).

The original heuristic evaluation have had its strengths explored (Nielsen & Molich, 1990; Jeffries, Miller, Wharton & Uyeda 1991); Iannella, 2009) and shortcomings (Jeffries & Desurvire, 1992; Nielsen, 1994a; Doubleday, Ryan, Springett & Sutcliffe, 1997; Kantner & Rosenbaum, 1997; Hertzum & Jacobsen, 2001; Holzinger, 2005; Ling & Salvendy, 2005) explored by many researchers and the common conclusion is that heuristic evaluation despite its challenges is an option worth utilizing. But it is uncertain whether these challenges and advantages are limited to them and can be applied to user experience heuristics.

1.1 Research question and method

According to the topic, this thesis will explore what advantages there may be for the evaluating party in using both usability user experience heuristics. This thesis will also study whether there are challenges in using those heuristics. The focus of study is set slightly in favor of UX heuristics due to the fact that there is much more research done about usability heuristics. This thesis acknowledges that using heuristic evaluation to assess systems user experience will not provide the evaluator with the full image of the user experience that the system may eventually invoke. However, this thesis will argue for using user experience heuristics are a part of a thorough user experience evaluation because of its ease of use and ability to discover user experience issues before and in the first stages of utilization.

This thesis will contribute the scientific community by providing a clearer vision of the advantages and challenges for heuristic evaluation of user experience. It will provide arguments to be used to explaining the criterion for either choosing or not choosing to utilize heuristic evaluation methodology. It will also point direction to additional research venues worth seeing into.

This thesis aims to find answers to the research questions by analyzing relevant existing body of literature and case study framework. The two cases used in this thesis have been done for a service – Facebook, for which already exists user experience (Väänänen-Vainio-Mattila & Wäljas, 2009) and usability (Hart et al., 2008) case studies. Facebook has been evaluated against usability and user experience heuristics in 2009 and 2013 using the heuristics by Nielsen and Molich (1990) and Arhippainen (2009). User experience research and heuristic evaluations are both rooted within the Human-Computer Interaction (HCI) field and interaction design (Hassenzahl & Tractinsky, 2006)

1.2 Structure of the thesis

This thesis will continue the following way. First in it will be presented relevant literature about its contents such as usability, user experience and heuristic evaluation for both usability and user experience. Case study consisting of heuristics evaluations for Facebook's usability and user experience will be analyzed in the Chapter three of this thesis. After the case studies follows the analysis of findings from the case study against the existing literature. The discussion about the literature challenges and advantages will be presented in Chapter five and this thesis will be concluded in Chapter six, in which will also be presented some ideas and pointers for future research.

2.   Background

Being the hot topic of the last and current decade (Hole & Williams, 2007; Vermeeren, Law, Roto, Obrist, Hoonhout & Väänänen-Vainio-Mattila, 2010) a lot of material can be found on all the aspects - usability, user experience, heuristic evaluation, of this thesis; usability, user experience and heuristic evaluation. This material will be presented in this section of the thesis and will later on be referred to.

2.1 Usability

User experience cannot truly be understood without understanding the notion of usability - thus this section will describe and define the term *usability* and its benefits. The term usability refers generally to the technical aspects of a system that can be measured and quantified and therefore improved (Bevan, 2009). The term itself was formed in the early 80's to replace the current term "user friendly" (Bevan, Kirakowski, & Maissel, 1991).

The International Organization for Standardization (ISO) has defined usability to be:

> *"The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"*

[ISO 9241-11:1998, definition 3.1]

Preece, Rogers and Sharp (2002, pp. 14) have identified six usability goals; effectiveness, efficiency, safety, utility, learnability and memorability. These goals play a vital role in ensuring the usability of a new system. Effectiveness is the degree which the system can help users achieving their goals. For example does the system allow users to access all the necessary data for them to complete their tasks? Efficiency means the least amount of work in order to complete the tasks the user desires. Safety is to prevent the users from entering into hazardous system states that may jeopardize user-created data or to prevent systems cause dangerous situations. Utility is the systems' ability to provide the user with the necessary tools to complete their goals. For example a text editor without the ability to write using keyboard would be a very low utility tool. Learnability refers to the systems' ability to be learned without watching tutorials or reading lengthy manuals. The basic functions needed to complete rudimentary tasks should be easy and fast to learn while advanced functions needed to complete more complex tasks can still take more time and effort to master. Memorability is the systems' ability to provide visual guidance and cues to help users remember what the goal of a function is.

Usability goals are a good tool to help create utilitarian systems with no regards to hedonic aspects - systems that are simply created to serve one purpose, finding a solution to a certain problem. Good example for such a system would be a text editor. Created to help users write, this tool is a perfect example of a utilitarian tool.

According to Holzinger (2005) usability design should be done early in the user interface design phase. The later it is done, the more costly and difficult the fixes will be

and the changes tend not be made (Boehm, 1988). The resistance most often comes from developers who feel that there are not usability issues to be fixed even if there are usability recommendations for the developed system.

Usability is evaluated in many ways, but the evaluation is not always for good. Usability evaluations in general have a shortcoming; they focus mainly on the negative aspects of the system, leaving unnoticed the good things the product has to offer. Another shortcoming is that through generally accepted norms about what good usability is - the more novel and innovative ideas may be disregarded to conform to the norm. (Greenberg & Buxton, 2008)

2.1.1 Consumer usability benefits

But are there benefits for good usability? Jokela (2004) has told three stories in which bad usability has either been devastating for a product's success or it has not had an effect on the customer acceptance what so ever. Jokela (2004) shares the usability into three groups; must-have, more-is-better and attractive usability. Through these three groups he explains that usability is not a static quality that either is or is not. The products may have usability issues but still strike through as innovative and good products.

Jokela (2004) states that users may not be neither content nor dissatisfied if the product they are using fills their must-have criteria. The must-have criteria for a phone might for example be able to be used to make phone calls and to send text messages. In modern phones many more attributes are expected and thus could be considered must-have qualities – such as the ability to connect and browse the Internet. But in order for customers to be happy, more is needed – more-is-better. In addition to the must-have attributes, more-is-better attributes are needed to make the users feel content. For example shortcuts, quick dial and possibly audio controlling for mobile phones are needed to make the customers, also known as users, content. Much-is-better attributes are necessary but alone do not count for good usability, lack of must-have features will lead to customer dissatisfaction.

The attractive-usability attributes come very close to actually being user experience attributes. Attractive-usability is something that is unexpected and unspoken needs of the customers. Such attributes might me highly customizable desktop which brings us to identification (self-expressing through an owned object). (Karapanos, Zimmerman, Forlizzi & Martens, 2009.) These attributes are needed to provoke amazement and awe towards the product and are thus needed for the ultimate usability experience and user satisfaction, which is one main factors of good usability (Jokela, 2004; Bevan 2009).

2.1.2 Business usability benefits

Good usability is clearly visible and in many cases expected for the end users. For businesses, both the creative and the using party, the benefits of usability are not as evident but they do exist and in time can lead to huge savings. Businesses in general are more interested in the applicability and efficiency of a product for specific problems. In general it could be said that if it works, do not fix it. This is not entirely true, usability issues may in fact lead to great expenditures in the long run. Jokela (2004) has stated that advantages for good usability for distributors and manufacturers as savings from not needing to invest as much in customer support and also as employee training savings.

Because of the intertwined and inseparable status of usability and user experience (Vermeeren et al., 2010), it can be stated that good user experience requires good usability and good user experience therefore also provides the benefits of good usability. Donahue (2001) has listed the benefits of usability on the manufacturing party; reduced development and maintenance costs, improved productivity and efficiency, reduced training costs, lower support costs, reduced documentation costs, litigation deterrence, increased e-commerce potential, competitive edge, advertising advantage and better notices in the media.

Reduced development and maintenance costs come from fixing usability problems in the development cycle, when it is cheaper. More savings come from avoiding arguing over decisions that are already covered by usability guidelines. Users will eventually user-test the program and delivering them a faulty program with usability issues may anger them and provide the company with bad reputation and loss of customers. Fixing the users' problems over phone also costs the manufacturer and distributor a lot of time and thus valuable working hours. (Donahue, 2001; Weinschenk 2005.) Weinschenk (2005) has supported Donahue's statements with examples of companies saving up to 90% of development costs by solving usability problems as early as in the design phase. She also claims that the cost of correcting usability issues after release is 100 times more costly than in design phase.

Improved productivity and efficiency is one of the greatest benefits of good usability. Even if the system works, inefficiency can cost the end users a great deal of money. An example given by Donahue (2001) about a company which improved a system's login feature's usability saved the company 41,700$ on the first day of use with the cost of only 20,700$ - the savings over the course of one year are easily induced. Weinschenk (2005) explains that users make fewer mistakes using a system that is developed using usability principles, thus saving time and money. Good usability also prevents stress. People tend to avoid using systems that cause stress and when forced to work with such systems, the productivity lowers considerably (Donahue, 2001; Weinschenk, 2005).

Reduced training costs are simply the result of a more usable system. If the users learn how to use the program easier and the program is easier to use in general, lesser training is needed. For every hour spent training is an hour off from work and an hour that the trainer needs to get paid for. (Donahue, 2001; Jokela, 2004; Weinschenk, 2005)

Lower support costs derive from users not needing to call the manufacturer or distributor for support. This saves time and money for users, manufacturers and distributor. The retailing party does not need to invest into as many help-desk workers. (Donahue, 2001; Weinschenk, 2005)

Reduced documentation costs derive from predictable and consistent interfaces. Being both predictable and consistent, they are easier to document. Good usability may even in some cases eliminate the need for manual which saves the manufacturer a lot of money not having to print and distribute it along with the software. (Donahue, 2001; Weinschenk, 2005)

Litigation deterrence; IT manufacturers may in some cases be sued for program faults. Though it's not as usual than in other manufacturing businesses the remaining risk is mitigated by using commonly agreed usability guidelines. (Donahue, 2001; Weinschenk, 2005).

Increased e-commerce potential is where usability has perhaps the most direct link with profitability. Nielsen (2003) has pointed out that people, informavores, hunt the information or the item they are looking for and only stop if they believe that they are getting closer. If the information or the item is wrong, or it is too hard to find, people move on to another site to seek what they are looking for. Forrester research (Manning, McCarthy, & Souza, 1998) states that hard-to-use websites drive away customers which forfeits revenue and erodes brands. Donahue (2001) states, what seems to be more like common sense that sites which are easier to use, aka more usable, are more enjoyable to use and users find what they seek easier.

Competitive edge is gained from providing the users with attributes they want. Consequently usability is something the users are wanting. Thus when users find a program or a service which is usable they feel they are respected and not taken for granted. Nowhere is the competitive edge as great as in e-commerce services, in which the users' attention span is very short and they leave as soon as they find out that the service is not up to their standards of usability. (Donahue, 2001.) One example of good usability is postponing the registration until the very latest possible stage of online shopping. Users want to build up trust with the service before they are willing to disclose any private information. (Nielsen, 1999; Donahue, 2001.) Nielsen (1999) points out that registration in the beginning of browsing is very similar to placing a security guard at the entrance of a store asking all possible shoppers to show ID's before entering the store.

Advertising advantage; good usability helps users to differentiate between different services. Good usability has been an important part of advertising from and before 2001 (Donahue, 2001). Many companies such as Microsoft, Compaq and many others have been advertising on usability as early as 1994 (Wiklund, 2012).

Better notices in the media. Usability is among the criterion through which product and services are evaluated. If the service is highly usable, it will be shown in the reports about the new products which can be considered a free advertisement and from a neutral party. (Donahue, 2001) The same applies to bad usability, when a program or a service is considered having low usability it will also be included in the reports. The latest incident where bad usability has been associated with a product was the launch of Jolla mobile phones, which received bad reviews because of the conceived lack of usability (Savov, 2013).

2.2 User experience

Usability research has been one of the leading areas in the field of Human Computer Interaction, but has lately been deemed too narrow a view to cover all the aspects of interaction and consumption of services (Law et al., 2009). As usability is the way to access the functionality of the system, it does not take into consideration whether the same functionality could be achieved otherwise more pleasurably (Battarbee & Koskinen, 2005). As Jordan (1998) has stated; usability does not take into consideration the positive feelings that a product should evoke. Products are not anymore just neatly packaged function but experiences which offer that functionality (Hassenzahl, 2005).

User experience as a term was conceived in the mid 90's and has ever since been a buzzword amongst the scientific community (Hassenzahl & Tractinsky, 2006; Ardito, Costabile, Lanzilotti, & Montinaro, 2007). Preece et al. (2002, pp 18) noted that with the emergence of new technologies has brought into attention new concerns. This notion lead them to state the user experience goals which are the following: satisfying, enjoyable, fun, entertaining, helpful, motivating, aesthetically pleasing, supportive of

creativity, rewarding and emotionally fulfilling. This is a step forward from the efficiency considered usability goals into goals that try to capture the *feel* which the interaction with the program of service provides.

The user experience goals and usability goals cannot all be wholly obtained within a single system. Many of those goals may be contradictory to one another. For example the system cannot achieve absolute security without compromising the ease of use. Thus Preece et al. (2002) note that user experience goals are contextual and the desired set of these goals depend on two factors; the system that is being designed and the users to whom the system is designed to. The key is to find balance between the two sets of goals which offer the user both best possible usability and the best possible user experience.

Many researchers agree that user experience cannot be separated from usability (Vermeeren et al., 2010), which instead is subsumed by it. Usability holds an intriguing notion of satisfaction. Satisfaction is not a typical aspect since there is no absolutely certain way to design for. Hassenzahl (2005) explains how satisfaction is actually the sum of hedonic and pragmatic qualities, which therefore also implicates a strong relationship between usability and user experience.

The user's perspective (Figure 1) shows that user experience is very situational, experiences may depend on the skills of the user, the context of use or ethnical backgrounds. In a nutshell the experience is very subjective (Hassenzahl, 2005). User experience still suffers from multitude of different definitions (Law et al., 2007; Hassenzahl, 2008; Appendix A) and assumptions.

User experience has been defined in ISO DIS 9241-210 (2010) as:

> *"a person's perceptions and responses resulting from the use or anticipated use of a product, system or service"*

Even though this definition can provide with a limited understanding about the concept of UX, it is a wider approach to HCI than what the definition of usability has previously offered. So what is meant by good user experience?

User experience is shared into two classes, hedonic experiences (Hassenzahl, 2007) - the experiences derived from i.e. aesthetics and enjoyment, and pragmatic experience – the experiences derived from usability and applicability. Hassenzahl (2007) explains the two different dimensions as different goals. Pragmatic experience refers to do-goals, being able to do things such as making a telephone call. Hedonic experience refers to be-goals, which able you to be something while completing the do-goals - for example being connected to others while for example listening to radio.
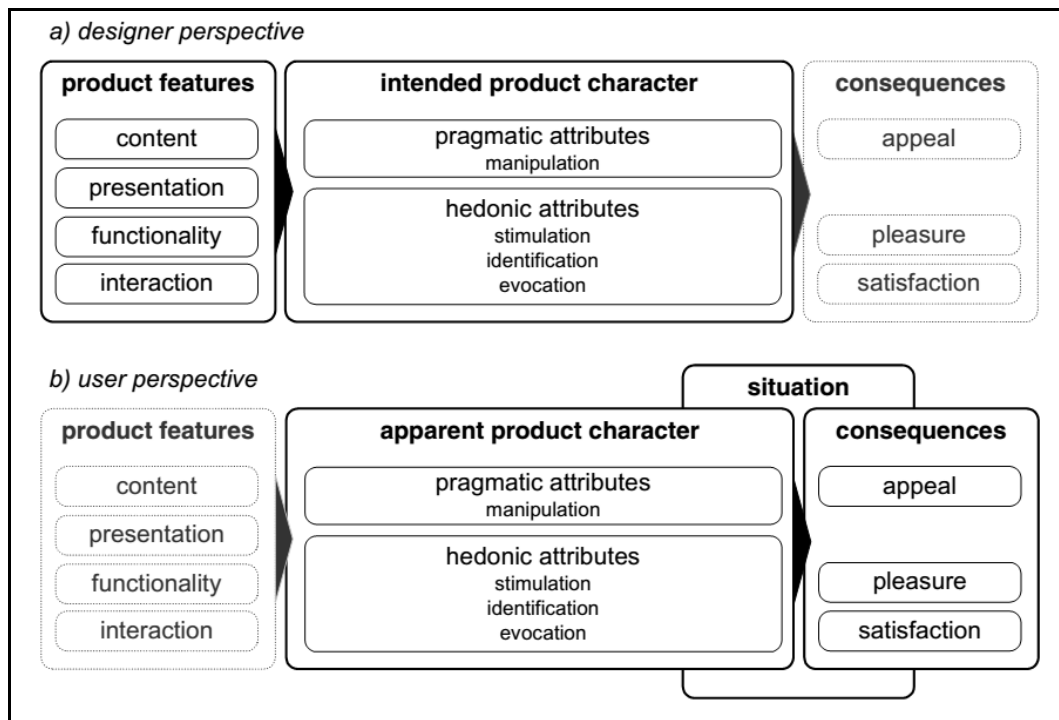
Figure 1. Key elements of the model of user experience from (a) a designer perspective and (b) a user perspective (Hassenzahl, 2005).

While practical goals are basically about manipulation, the user needs something done and a tool to do it, the hedonic attributes can be shared into three categories; stimulation, identification and evocation. Stimulation is closely related to flow (Csikszentmihalyi, 1988) and is about the service or product being able to stimulate the users by providing impressions, opportunities and insights. For example functions that are not used but which are available can be considered hedonic. They still may help with goal fulfilment, aka practical goals, by motivating or offering a possible new way to achieve the goal. (Hassenzahl, 2005.)

Identification is an important aspect of user experience. Users want to derive value from the services and products they are using and construction of identities is one way to co-create value. (Tuunanen, Myers & Cassab, 2010.) People want to communicate their identities to others, most often in advantageous way. Thus a product with the ability for self-identification is conceived to have a better user experience. The final aspect of user experience is evocation, which stands for memory provoking. An old game may provoke the sensation of nostalgia even when the game itself may be outdated. While the game itself has nothing else to offer, the feeling itself is strong enough reason to compel people to play it over and over again. (Hassenzahl, 2005.)

Hassenzahl (2008) and Deng, Turner, Gehling, and Prince (2010) argue that the actual use of technology itself can be a source of pleasure if it manages to exceed the users expectations and provide them more than mere task completion this notion is also supported by many other researchers. (Hassenzahl, 2007; Ardito et al., 2007.) Hassenzahl (2005) has also pointed out that commitment to a product with pure practical qualities is delicate at best and can easily be severed if a better option to fill the same task is available - whereas commitment to a product with only hedonic qualities can be very strong and hard to replace.

User experience is very situational, same qualities may evoke a feeling of good experience at some situation but in other situations the feeling evoked can be closer to

frustration. A software which is designed easy to use and learn at the cost of additional functionality may feel innovative and pleasant to use at first but eventually the loss of additional functionality may become to feel like an encumbrance. Karapanos, Zimmerman, Forlizzi and Martens (2009) have studied the long term user experience and come to the conclusion that while initial acceptance has been studied to some detail, research about long term acceptance is still lacking.

2.2.1 Initial user experience vs. long term experience

Initial feeling of good experience may change over time and in some cases turn into frustration. Koca, Funk, Karapanos, Rozinat, van der Aalst, Corporaal & Brombacher, (2009) have acknowledged that the amount of fully functional returned products has risen to 50% in the 20th century – from the near zero level in 1980's. This is due to change in people's attitudes toward products and services. Most of these cases were about the product not fulfilling the expectation of the users or not offering good enough experiences.

According to Karapanos et al. (2009) there are three forces; *familiarity, functional dependency* and *emotional attachment*, that affect long term acceptance and the transition of users experience between acceptance's three stages; *orientation, incorporation* and *identification* (Figure 2).
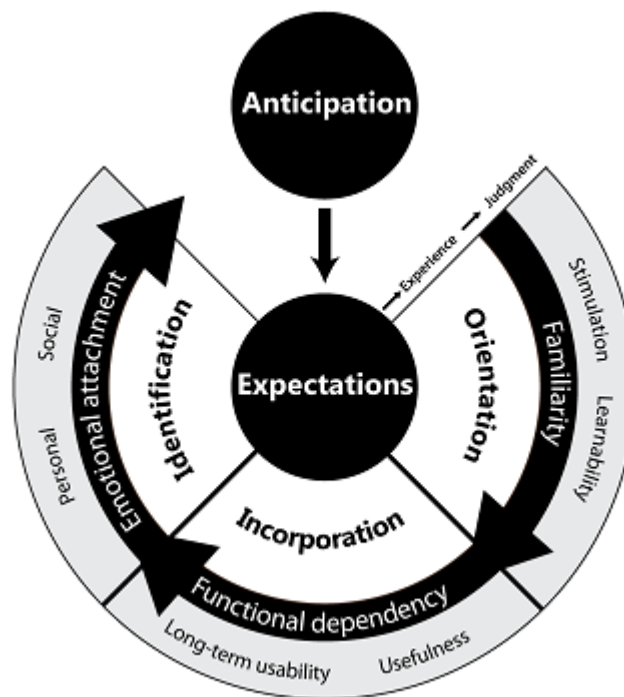
Figure 2. Temporality of experience (Karapanos et al., 2009.)

The definition of user experience (ISO DIS 9241-210, 2010) holds a notion of anticipation as a factor that affects the user experience. Karapanos et al. (2009) explain anticipation as being the act of anticipating an experience which results in the formation of expectations. *Anticipation* happens before any actual use of the product or service and can be either positive (*opportunities for positive experiences*) or negative (*fears of negative implications*).

Orientation is the stage of initial feelings from when a product or a service is taken into use. This stage includes the feelings of awe and frustration that stem from either novel

features or learnability flaws. *Familiarity* was the most prominent factor in orientation phase including aspects such as *stimulation* consisting of *visual aesthetics* and *aesthetics in interaction,* and *learnability*. The feelings that stem within the orientation stage fade very quickly, in the case of iPhone orientation, it only took a week. (Karapanos et al., 2009.)

Incorporation is the stage of product or service becoming meaningful in the users' daily lives. This stage is dominated by the products usefulness and factors such as learnability lose their relative importance. In incorporation stage two aspects are prominent in determining user experience; *long term usability* and *usefulness*. The latter can be shared then to *fast access to information, alleviating boredom in idle periods, capturing momentary information* and *avoiding negative social situations.* There is the possibility that the product or service may cause bad experiences by provoking the user with *long term usability issues* or *usefulness problems.* (Karapanos et al., 2009.)

The final stage; *identification* is the stage when a product either connects the user to a community or differentiates the user from others. This is done by product participating in the users' daily interactions and promoting their identity in these interactions. Identification includes two perspective; *personal* and *social. Personal* experiences perspective includes; *personalization* and *daily rituals,* whereas the social perspective includes aspects such as; *enabling self-expression* and *creating a sense of community*. Enabling self-expression is a way for users to differentiate themselves from others whereas the latter is a way for the users to feel part of a community. (Karapanos et al., 2009.)

2.3   Usability heuristic evaluation

The origins of heuristic evaluation span over a long period of time and topics ranging from philosophy to HCI. The first set of heuristic for HCI, created by Jacob Nielsen and Rolf Molich (1990), were however only just published in 1990. The reason behind the creation of the heuristics was that before their creation, thousands of guidelines for user interface creation existed. Using these guidelines, even to find the right one to use could take a very long time. Nielsen and Molich (1990) created the heuristics that would be easier to learn and faster to use. The heuristics received lot of attention and quickly became one of the main methods in assessing usability of a user interface. (Rosenbaum, Rohn & Humburg, 2000; Holzinger, 2005; Ling & Salvendy, 2005.)

Nielsen was not fully pleased with the original heuristics and in 1994 created a revised version of these heuristics. The new revised heuristics, which are still in use today, include the following (Nielsen, 1994a); visibility of system status, match between system and the real world, user control and freedom, consistency and standards, error prevention, recognition rather than recall, flexibility and efficiency of use, aesthetic and minimalist design, help users recognize, diagnose and recover from errors, help and documentation.

1. Visibility of system status means that users need feedback from their actions in a timely and appropriate manner.

2. Match between system and the real world states that the system has to use known terminology and show the information in a manner both familiar and logical for the user.

3. User control and freedom means that the user must have a way to escape from any system errors. This also means supporting undo and redo.

4. Consistency and standards mean that the user is aware of the meaning of different words, situations and actions.

5. Error prevention is even better than error handling. This means either eliminating the possibility of errors or at least asking the users' confirmation before committing to an action that is prone to cause errors.

6. Recognition rather than recall, stresses the importance system visibility and easing the memory load of the user. Different options for users should be visible and system instructions easily retrievable if needed.

7. Flexibility and efficiency of use means that users should be able to use and create accelerators for common and frequent actions.

8. Aesthetic and minimalist design stands for easing the cognitive overload for users. Only the needed and relevant information should be shown at any time and in an aesthetic fashion.

9. Help users recognize, diagnose and recover from errors means that error messages need to be plain, state the problem and offer a solution.

10. Help and documentation should be available for the user at any time.

Heuristic evaluation, sometimes referred as a discount method, gained worldwide acknowledgement swiftly and became well known and one of the main methods for usability testing (Ling & Salvendy, 2005). This is due particularly to two reasons; the two main reasons for its success are the relative ease and cheap cost, for example when compared to traditional user testing, of using the heuristics. (Nielsen & Molich, 1990; Iannella, 2009).

There is also a third dimension which partly adds to the success of heuristic evaluation – time. The method is also quite fast compared to other discount methods. The basic idea of the method is to have a couple of usability experts who evaluate a system or product individually against the given heuristics and note the deviations from the heuristics. (Ling & Salvendy, 2005.) One aspect which also makes heuristic evaluation such a success is the fact that it can be applied early in the design phase and find issues even before the development has begun and when the repairs are a lot cheaper (Nielsen, 1992). None the less, for heuristic evaluation to be successful there needs to be at least a couple of evaluators.

The reason for the need of couple of evaluators is explained by Hertzum and Jacobsen (2001) who state that the overlap of findings between two evaluators ranges from 5% to 65%. But this problem is easily resolvable by having at least three evaluators - Nielsen (1994a) notes that five evaluators is the optimal amount. Five evaluators can find around two thirds of the usability issues and has the highest benefit-to-cost ratio. Research has been done which indicates that heuristics are also a better way to find specific usability problems than for example cognitive walkthroughs (Desurvire, Kondziela, & Atwood, 1992). Other researchers have also pointed out that heuristic evaluation found most serious problems with the least amount of work compared to other methods (Jeffries, Miller, Wharton, & Uyeda, 1991).

2.3.1 Challenges of usability heuristic evaluation

As stated earlier, one of the minor drawbacks of the heuristic evaluation method is that it requires at least a couple of evaluators – three to five would be optimal. This is due to the fact that different people find different usability problems and rate them differently based on their severity. (Nielsen, 1994; Hertzum & Jacobsen, 2001) Even though this is not a major obstacle it something to keep in mind if there are not many usability experts available.

Another problem of heuristic evaluation is the reliance on the sole expertise of the evaluators assessing whether any of the observed faults actually violate the heuristics which are merely guidelines (Doubleday et al., 1997; Ling & Salvendy, 2005). There is always a chance that an unexperienced evaluator may give grave error a pass because he or she did not consider it a deviation from the heuristics. This is also a case where having multiple evaluators helps, if one discards a problem, the other may find it. Holzinger (2005) points out that the expertise required to conduct heuristic evaluation is not as high as in other methods. However, this adds to the uncertainty of thoroughness and reliability of the found usability issues. Holzinger (2005) also points out that in some cases it can be appropriate to use less-experienced evaluators depending on who is available. Heuristic evaluation done with less-experienced evaluators is still a viable tool to find usability faults in high degree.

Another minor problem is that heuristic evaluation tends to find more false-positive errors – meaning errors that won't occur with actual end users and thus need no fixing. The only real drawback with this problem is that a high amount of false positives may result in real problems being ignored. Heuristic evaluation does not take into account the intelligence of the users, the users being able, through logic; make logically possible problems empirically impossible. (Cockton, Woolrych, Lavery, Sears, Jacko, Tsuchiya, & Grandy, 2008.)

Heuristic evaluation is reaching for the situation which Nielsen wanted to correct in early 90's. With so many different kinds of services available and the original heuristics not being able to cover these new kinds of services and products, the amount of different variation of heuristic evaluation methods is rising. As Holzinger (2005) has explained, the selection of the right set of heuristics is vital for the success of heuristic evaluation – especially the additional heuristics for web-based services are critical. The problem becomes the selection process from a large pool of possible set of heuristics.

Yet another drawback is the separation from actual end users (Kantner & Rosenbaum, 1997). This issue can be considered both strength and a weakness. Because with heuristic evaluation there is no need to use the actual end users of the product, the evaluation process can be done very rapidly and without the trouble of seeking viable test users. However, without utilizing the end users, some of the unexpected issues may remain unnoticed. Heuristic evaluation does not take into consideration unknown needs that might emerge from user testing. (Holzinger, 2005)

The final major problem identified in an article by Ling and Salvendy (2005) is the heuristic evaluations inability to offer any support for countering the found issues. The depth in which heuristic evaluation offers recommendations for fixing the found issues is solely based on the expertise of the evaluators. This can especially be an issue when there are no experts available and the evaluation process of being conducted by less-experienced evaluators.

2.3.2 Heuristic evaluation for user experience

Even though there are several shortcomings, heuristic evaluation is still the main *"discount"* (Nielsen, 1989; Jeffries et al., 1992) method for evaluating user interfaces (Rosenbaum et al., 2000; Holzinger, 2005; Ling & Salvendy, 2005). With such a success at usability evaluation field, could heuristic evaluation help assess and evaluate the user experience with such a low cost and high efficiency?

Heuristic evaluation has mainly been a tool for assessing service usability but has also been adapted for user experience. There are 10 user experience heuristics; Ensure usability, Provide utility matching with the user's values, Surpass the user's expectations and minimize the gap between negative expectations and real usage, Respect the user, Design the product or service to fit the intended contexts, Provide several ways to interact and leave choice for the user, Respect user's privacy and security, Support user's activities but do not force them, Go for a perfect visual design and Give a surprise gift. (Arhippainen, 2009. pp. 223-224.)

1. *Ensure usability* is a vital step for ensuring good user experience. Users experience the usability while using the product or a service, which is why it is important to design the product via usability guidelines.

2. *Provide utility matching with the user's values*. Users will perceive the products utility. A useful product may have some of its shortcomings overlooked. Usefulness is closely related to users' personal values and for this reason these two should coerce near equilibrium.

3. *Surpass the user's expectations and minimize the gap between negative expectations and real usage*. People tend to feel negative emotions towards new products; this may be because of previous use of similar products or via rumors. Catch people's attention and when they start using the product, exceed their expectations by being easy to use, pleasant or by offering utility to the users.

4. *Respect the user* (know the needs of the users). The product should take into consideration users' values, skills and needs. The product should be made keeping the users in mind; and thus it has a better chance to offer the users better user experiences. This could mean involving the users in designing of the product.

5. *Design the product or service to fit the intended contexts* means that the designer needs to take into consideration the circumstances in which the product will eventually be used. Some products are used alone, some together. Some are used in harsh environments and in a certain cultural context. These circumstances all affect the user experiences one way or another.

6. *Provide several ways to interact and leave choice for the user*. Different people like to use products differently, thus it is important to allow them to do so when possible. For example by offering manual option to an automatic one or a mouse and keypad as an option to motion guidance.

7. *Respect user's privacy and security* means that even when technology has made peoples' attitudes more open towards gathering and sharing of information, they are still concerned about their personal privacy and security. Uncertainty on the products personal security and privacy has a negative impact on user experience.

8. *Support user's activities but do not force them*. Product should be designed in a matter that supports user in what he or she is intended on doing. The product should not enforce the user to compromise.

9. *Go for a perfect visual design.* The aspect of visual design has two meanings. While good visual design can make the service or product easier to use by offering more consistency, understandability and coherence – the other meaning is aesthetics. Aesthetics can make the service more pleasurable to use for example by using generally accepted visual aspects and colors that promote positive emotions.

10. *Give a surprise gift*. Good usability is not enough, steady feeling of good experience is not enough. Design something extra for the users to surprise them positively, to make them feel that the product is alive and increase the good user experiences and makes the users happy.

The usage of the above heuristics is tutored in the more recent article by Arhippainen (2013). The use of these heuristics is not limited to any specific products or services but instead aim to be general guidelines that can be applied to any situation. There are also specific heuristics that should better take into consideration the subtle nuances of different types of products. Example of such specified heuristics are the seven *web service UX heuristics.* (Väänänen-Vainio-Mattila & Wäljas, 2009)

1. *Usage and creation of composite services.* Users may add or create additional service components to the current service.

2. *Cross-Platform service access.* Users are able to use the service with different platforms, such as smartphones and PCs.

3. *Social interaction and navigation.* Users can interact socially via the service and use other users' navigation history while interacting with the service.

4. *Dynamic service features.* Users see dynamic changes in the service UI.

5. *Context-aware services and contextually enriched content.* Depending on the context of use, the service offers the user different meaningful content.

6. *General UX-related issues.* These can be considered to be a link to the 10 heuristics proposed by Arhippainen. (2009)

7. *Findings outside heuristics 1-6*

However, assessing UX is not as simple as assessing usability. Even if the heuristic evaluation for user experience is procedurally similar to evaluating usability, there are fundamental differences between usability and user experience. These differences challenge the applicability of UX heuristics and are explained in more detail in the discussion section.

3.   Two case studies of Facebook

This section will consist of two case studies for a heuristic evaluation of Facebook - a service which has a very large user base and has been around since 2004. The first case study consists of usability and user experience evaluations done for Facebook in 2009 as a part of a university course practical work. At the time the evaluation was conducted by four evaluators. At the time Facebook had many serious flaws in its usability, this thesis tries to find out whether the ever growing popularity has managed to smooth the flaws.

The second case study consists of usability and user experience evaluations done for Facebook in 2013. This time there was only one evaluator doing the evaluations. The reason why both of these evaluations are considered in this thesis is that with additional data, the results are more reliable and more reliable conclusions can be drawn about the effects of the evaluators' expertise.

The method of evaluating both usability and user experience of Facebook were methodologically similar in years 2009 and 2013. In 2009 the team had two similar evaluation models through which the notes on found deviations were recorded. The models were created by Arhippainen, but have only just been published in 2013 (Appendix B). The team used the service while having copies of the heuristics with them. The method was used so that the team could see the heuristics and easier find the possible deviations. In 2013 the same method was used and due to the method, documentation from 2009 is comprehensive enough to use the results in this thesis.

This Chapter will follow accordingly; first a brief summary of Facebook is told followed by a heuristic evaluation of Facebook's usability from the year 2013 and the summary of found deviations. The section 3.3 will present a heuristic evaluation of Facebook's user experience and summary of found deviations - also from 2013. The third section consists of heuristic evaluations done in 2009 with four evaluators and the summary of found deviations. The final section 3.5 will combine and sum up the findings from the case studies.

3.1 Facebook

Facebook has in the last five years become one of the greatest success stories of web 2.0. With its user base of 1,18 billion users, as of June 2013 (https://newsroom.fb.com/Key-Facts) it is also one of the largest online services available for advertising purposes as well as keeping in touch with acquaintances and friends. Facebook is available in almost all the countries of the world and often countries who do not allow direct access to it have had their own versions of it made.

The basic working principle of Facebook is that once user logs in, he or she can find most her friends from Facebook. They can be added as friends and sent messages to. Users can upload pictures and share them with friends or post them as public for everyone to see. Users can mark friends to uploaded pictures or posts thus referring this particular friend in the post. For many users, Facebook is a place where they can keep tabs on what is going on in their friends and acquaintances' lives. Thus the service is not

meant to achieve any specific goals, perhaps expect the users being able to share what they will with the people they will.

Facebook is also a service which offers its users many activities such as playing games, either alone or with friends. The users can also create content that other users can use and browse. The content might be free or not, thus the content creators can actually earn by making content that attracts users. Even with free content the income can be based on advertising. Even though this is an option that Facebook offers, it will not be taken into consideration in the heuristic evaluations done in this thesis.

But what makes Facebook so popular? Is it because it is so good to use or because it offers a very good user experience to the users? As stated earlier, if usability was the only factor concerning system acceptance and popularity, Facebook should have failed, at least according to Hart et al. (2008) who in their research found out that Facebook had failed eight out of ten usability heuristics which indicates that it should be somewhat unusable.

In this thesis Facebook will be tested for both usability and user experience aspects. It will be interesting to see how the usability has changed in the five years it has been since Hart et al. (2008) have analyzed it last – excluding the evaluations done in 2009.

3.2 Usability heuristic evaluation 2013

Nielsen and Molich (1990) have created the original heuristics on usability evaluation. Facebook has been evaluated according to these heuristics by Hart et al. (2008). At the time it failed eight out of ten heuristics. The heuristics 4, 5 and 6 (consistency and standards, error prevention and recognition rather than recall) in particular had problems.

Heuristic evaluation will be done as follows: Under each heuristic will be written findings concerning main functions and actions available to users. These findings will be general and descriptive in nature. After all the heuristics have been assessed against Facebook, the findings will be summarized in the next section.

Heuristic evaluation of Facebook's usability will be done in this section. It is expected that problems about usability will be found but this thesis will not suggest any fixes for those problems and merely uses the results of this section in assessing whether heuristic evaluation is a viable way to assess user experience of a web service. The usability heuristics are in their original form as presented by Nielsen and Molich (1990).

3.2.1 Exploration of flaws through heuristics

1. *Visibility of system status means that users need feedback from their actions in a timely and appropriate manner.*

   Facebook may have many usability issues but this is not one of them. Feedback is almost imminent after an action has been completed. For example when writing posts, the update is done instantly and the user can see his or her post appear immediately. When pointing over other users' profile picture the user is given a short review of their friendship; mutual friends, the possibility to open a friendship page and the possibility to write a message to the friend. The short delay before showing the information is also a nice aspect because it prevents information overload while browsing

There are also other aspects that have been taken into consideration with Facebook. When trying to use application in secure https, the applications without support for https inform the user about it. And when creating an account, the chosen password strength is updated accordingly to given password. Also while uploading pictures or videos a progress bar is shown presenting the status of the upload.

However users can quickly become lost in the maze of different pages that is Facebook. Although there are few ways that users can go back to the main page, the different pages are not always titled which can make the user confused about the page he is in.

2. *Match between system and the real world states that the system has to use known terminology and show the information in a manner both familiar and logical for the user.*

   This heuristic begins to unravel some problems in the Facebook design. First problem is the use of non-standard icons; globe icon for notifications when common sense might suggest using exclamation mark, a common icon meaning notification. Messages are retrieved via function with a conversation bubbles icon – although Facebook does not discriminate between quick messages and other messages the common standard icon for messages is an envelope.

3. *User control and freedom means that the user must have a way to escape from any system errors. This also means supporting undo and redo*

   In Facebook it is possible to make mistakes, post a comment or create a new post that you did not mean to. There is also a way to delete or edit your comment. However, as the heuristic states, there should be a clearly marked "emergency exit" and a support for undo and redo, which does not apply to Facebook in full. The exit is hidden unless you know where to point your mouse to and even then, the delete / edit button is hardly visible. When trying to delete a posted comment on Facebook wall, user is first required to find the delete / edit button, from which, by pressing, emerges a drop down menu from which the user presses delete, and after that the system still requires the user to confirm to deleting the post. Facebook does not support undo or redo.

   There is no clear way to quit apps except the Facebook logo on the left corner, which cannot be considered clear. This may not be true for all the applications, but at a quick glance at them, it seems to be the case.

4. *Consistency and standards mean that the user is aware of the meaning of different words, situations and actions.*

   While browsing someone else's wall, the icons for posting and update status have the same image as in your front page but different textual explanation. However their function is almost similar so this can perhaps be overlooked. Same bodes for uploading photos to someone else's wall.

   What is strange is that the user has more options available for posting after they have navigated into their own page. The front page of Facebook does not have all the posting options visible but is exactly the same function as in the user's personal page.

5. *Error prevention is even better than error handling. This means either eliminating the possibility of errors or at least asking the users' confirmation before committing to an action that is prone to cause errors.*

While the possibility to make human errors in Facebook is present, the conditions for errors have been lessened by the inability to post an empty post. This bodes also for the additional information. Even if you can just post your location to the wall, you cannot post that you're just reading, at least with the "activity" button.

Double posting is prevented by a simple close function. Once a message has been sent, you need to write a new one, thus disabling the option to post the same message twice.

When searching for people, Facebook suggests the user a person based on the number of mutual friends. This makes the search faster and a bit more secure, if the number of mutual friends is high enough, the user can be more certain that it is the person he or she is looking for, thus preventing the friending of a stranger.

The "unfriend" function however does not require any confirmation at all. There is no way to revert this without sending a friendship request to the person. The unfriending is also invisible to the target that's being unfriended, which avoids an uncomfortable situation of explaining the unfriending.

6. *Recognition rather than recall, stresses the importance system visibility and easing the memory load of the user. Different options for users should be visible and system instructions easily retrievable if needed.*

Facebook supports this heuristic by showing users the available options visually in every situation. System instructions are however nowhere to be found. User has to actually visit a third party page in order to find the instructions on the basic functions of Facebook. This may not be the most crucial issue for an IT oriented user who will most likely learn the basics by simply testing. For inexperienced IT user, for example the elderly the lack of using instructions may prove a critical factor in determining between abandonment and continual usage.

7. *Flexibility and efficiency of use means that users should be able to use and create accelerators for common and frequent actions.*

Facebook does not support many accelerators but there exists some functions that can be considered accelerators. These accelerators include functions such as the ability to customize messaging groups. This function accelerates the user's ability to message to certain persons and groups without the effort of selecting the different persons all over again – the same applies for the creation of groups.

When searching for friends or groups, Facebook first tries to suggest to you your friends beginning with the given letters. This makes it faster to find a friend that you're looking for as the user needs in most cases only type the few first letters of his or her name. After if no results can be found within the users friends, it suggest a person based on mutual friends. This can make the search faster but it is also a great minor security feature.

8. *Aesthetic and minimalist design stands for easing the cognitive overload for users. Only the needed and relevant information should be shown at any time and in an aesthetic fashion.*

Facebook cannot in any way be described minimalistic in design. The whole idea is that your Facebook wall is filled with information that is posted by your friends. In addition to the posts, the user is also presented with information about who of their friends is online at the moment and a list of their friends' recent activities. However, one could argue that Facebook does this as aesthetically as possible.

Aesthetics are a hard topic and what is aesthetic to someone else, may not be the same for others. However in Facebook this problem has been solved by using only small amount of aesthetic elements. Different pieces of information have clearly been separated from other pieces, this dividing of elements makes the flood of information easier to divulge.

9. *Help users recognize, diagnose and recover from errors means that error messages need to be plain, state the problem and offer a solution.*

No errors were encountered during the evaluation period.

10. *Help and documentation should be available for the user at any time.*

Facebook has a help function from which the users can find help with the common problems which may occur while using Facebook. Help is based on search in which users type a word describing their problem. The search then suggests possible problem scenarios from which the user can choose a suiting scenario. If there is no such scenario, the user can specify the problem and thus find the scenario he is after. The front page also offers some basic topics that user can look for an answer.

If the user is not able to find answer to his or her problem despite best efforts, Facebook also has an option for reporting issues and seeking help from a Forum. This ensures that solution is found to almost any imaginable problem.

The help button itself is hidden under the settings button in the upper left corner of Facebook user interface. It is easily accessible but not accidentally, so the users do not have to worry about accessing help unintentionally or sending help requests accidentally.

3.2.2 Summarized findings (2013)

In this section, a summary of findings will be presented. The findings will be collected in a concluding table (Table 2), from which the findings are easier to quantify and present. The findings are also ranked from 0 to 4 according to their severity (Table 1). The principles behind the creation of the explanations for rating scales are procured from Nielsen's (1995) article *Severity rankings for usability problems.* There are three elements that form the overall severity; *frequency, impact* and *persistence.*

The *frequency* refers to the number of times the problem occurs within a given timeframe. If the problem occurs highly infrequently it is considered a mere nuisance and depending on the impact and persistence of the problem, the severity is then evaluated. *Impact* refers to the impact the problem has on the utilization of the service. When a problem occurs, will the users be able to deal with it or does it effectively prevent utilizing the service. Even the most crippling problems may not be the most severe ones if they happen highly infrequently and are easy to get rid of. *Persistence* is the problems attribute to pop up even if the user has dealt with it before. This can cause

annoyance and with the case of crippling problems, prevent the users from utilizing the program or service altogether.

Table 1. Rating scale severity rankings for usability problems

Severity of findings	
Rating	**Explanation**
0	Violates a heuristic but is not a usability problem
1	Violates a heuristic and is a superficial usability problem. The problem can be ignored or it happens highly infrequently.
2	Violates a heuristic and is a slight usability problem. The problem occurs more frequently and causes inconvenience to the users.
3	Violates a heuristic and is a major usability problem The service is still usable but causes inconvenience frequently thus frustrating the users.
4	Violates a heuristic and through inability to use may prevent utilization

Thirteen deviations from the heuristics were found while exploring Facebook. The number of deviations might be higher if the exploring would have been even more thorough. Time limitations and number of evaluators were a strong reason to keep the evaluation process short.

Table 2. Summary of usability findings

#	Problem	Severity ranking	Heuristic number	Broad Heuristic
1	Globe icon for notifications	0	2	Match between system and the real world
2	Text bubble icons for messages	0	2	Match between system and the real world
3	No titles on all of the pages	1	1	Visibility of system status
4	No clearly marked emergency exits for unintentional posts	2	3	User control and freedom
5	No undo or redo	3	3	User control and freedom
6	Three steps to remove a post	0	7	Flexibility and efficiency of use
7	No confirmation required for the unfriend function	3	5	Error prevention
8	Same icons but different explanations for posting options on friends' page	1	4	Consistency and standards
9	All the posting options not visible in the main page	1	4	Consistency and standards
10	No use instructions available	2	6	Recognition rather than recall
11	No accelerators for accessing posting or messaging functions	2	7	Flexibility and efficiency of use
12	Too much information shown in the front page	2	8	Aesthetic and minimalistic design
13	Help button hidden under settings button	1	10	Help and documentation

As expected, some usability issues were discovered in Facebook's evaluation. There were only few severe issues which was to be expected because the system has been around for almost a decade. Continuous feedback from the users is one of the main factors in reducing the amount of crippling usability issues.

There were two issues that were rated three in the severity scale; *No undo or redo* and *No confirmation required for the unfriend function*. Especially the first deviation is a severe one, because for example three steps are required to delete a post. With millions of users posting tens of millions status updates daily, from which a certain amount includes errors which force users to delete the post, an *undo* function seems critical for improving the usability of the service.

Second deviation ranked three in the severity scale is *no confirmation required for the unfriend function* – in general heuristic about error prevention. Even though unfriending is not a common function in Facebook, it can still happen. And when there is a possibility for error, with millions of users, they will most likely occur. This would easily be fixed by adding confirmation for such an action.

With Facebook being an integral part of many peoples' lives, issues ranked zero in the severity scales are often taken for granted and not seen as usability issues at all. Such is the case for example with icons. Users over the time have become accustomed the globe icon denoting notification even though the norm for denoting notifications would be an exclamation mark.

3.3 User experience heuristic evaluation 2013

The user experience heuristics used for evaluating Facebook's user experience were created by Arhippainen (2009). There have not been any previous published evaluations of Facebook using these heuristics. The evaluation will follow the previously done heuristic evaluation of Facebook's usability; first the service will be explored through user experience heuristics and then summarized in a table for easier quantifying and analysis. The heuristics are in their original form as presented by Arhippainen (2009).

3.3.1 Exploration of flaws through heuristics

1. *Ensure Usability. Users experience usability. Therefore it is important to ensure that the designed service or product is usable. Ensure usability by utilizing design and evaluation guidelines and practices of usability engineering.*

   Usability evaluation is done in heuristic evaluation of usability section. For redundancy reasons the results are not duplicated here.

2. *Provide utility matching with the user's values. Utility of the product or service effects on user experience. Perceived utility forgive lacks in usability or other product qualities. Utility goes hand in hand with user's values. The user balances between his values and the utility of the product when choosing to use it.*

   Utility and Facebook are not an ideal match. Because a mostly hedonic service with only a few utilitarian aspects it is hard to be evaluated for its utility. However, if the users' only interest in using Facebook is to connect to his or her friends and be able to send them messages, pictures, videos among other possible messaging forms, the utility is present. Facebook is the ideal messaging

center for informal communication which offers all the forms of messaging from video conferencing to email-type messaging. User's values are always independent of other users' but can be influenced by them. If the user values privacy, security and clarity on both aspects, the values may clash with utility.

The interface remains constant no matter where the user goes, the top bar allows users to quickly return to main page and quick messaging is always available. This creates additional utility since the users can browse at will and even occupy themselves with creative pastimes with all the while being able to connect with their friends.

Perhaps one aspect of utility where Facebook shines the most is the possibility to use its credentials to sign in into other services online. Using already created account to connect into other services save users time and eases frustration with not having to deal with many account creations. This of course in turn adds to switching costs for Facebook. The more services are intertwined with Facebook, the harder it is for users to stop using it.

3. *Surpass the user's expectations and minimize the gap between negative expectations and real usage. Often the user's expectations are negative for no reason. For example, expectations have been formulated via prior experiences or rumor of the product, and thus expectation may have nothing to do with the product in question. Therefore, the product should be able to catch user's attention in a positive way and get a user start to use the product, and then surpass his expectations by easiness, pleasure, utility, whatever quality could fit in the case.*

Using Facebook can be confusing at first. For first time users there is no tutorial available and users have to learn the basics themselves. In many cases it takes some time to learn the basics and more time to master the service. Meanwhile many users are being frightened by stories that Facebook shares all your images publicly online and the users have no control over what is shown and to whom. This fear is partly based on truth; the privacy settings in Facebook are still messy.

When a user adds a friend, he has to choose to what group to put the new user in. If all of the user's contacts have been assigned into groups, the groups can be used as a filter to who the messages and pictures are visible. But there is still a lot that can go wrong - when a user is posting pictures he has to select the privacy settings for that picture. If the user has added anyone of his friends to a wrong group it could mean there is potentially a risk of someone seeing the picture unintentionally. And even if the user manages to get it right, the settings must be changed back for the next upload. So there is a lot of confusion, at least for beginning or non-technology oriented users.

However, the utilitarian aspect of Facebook manages to uphold and even surpass expectations. Messaging itself is done in a rather efficient and easy-to-learn manner. The sheer amount of users in Facebook make it possible to find almost any person imaginable, once friended or even unfriended, there are messaging options available. In this aspect the users' expectations are most likely surpassed by the ability to find almost anyone and be able to socialize with them all in one place.

4. *Respect the user. Know the target user groups. A user's background has a strong impact on how he will perceive the product or system. In addition to the user's needs and actions, designers are required to understand the user's values, prior experiences, user type, skills, restrictions, etc. The better the service fits the user's world, the better experiences the user will have. Respect a user has a large meaning here. It can also refer to taking end-users into the design processes (participatory design).*

   With Facebook, this heuristic is almost inapplicable. With the amount of users as high as 1.18 billion (https://newsroom.fb.com/Key-Facts) it is very hard to try and know the different user groups. And even if the idea is to please the largest user groups, even the minorities are counted in millions and there is a lot of reason to cater for them too. At this point in Facebook service lifecycle it is hard to know all the user groups that are using the service.

5. *Design the product or service to fit the intended contexts. The service or product is always used in particular contextual circumstances: the user is using a product in a certain usage situation, in certain physical place with a specific company or alone through the certain cultural habits and way of life in a certain temporal moment. All these context factors have more or less impact on user experience.*

   Facebook supports a plethora of different contexts. This is the advantage of it being a web based service, it easily adapts to different contexts. There are several ways to use Facebook. It offers the users possibilities to use the service via personal computers, mobile phones and other handheld devices. As long as the user has an access to the Internet, Facebook is available. Users can use Facebook as a group chat service simultaneously or a person is able to post group messages and wait for the rest of the group's responses. Two users can have a video chat while uploading pictures and videos. User commuting to a work can browse Facebook in a car, train, bus via mobile phone and even post comments if he or she wants. As a web service, Facebook offers multi-contextual services in almost any circumstances - as long as there is an Internet connection available.

6. *Provide several ways to interact and leave choice for the user. We are different and we prefer different ways to act with products and services. Therefore, it is important to provide different ways to interact with service or product. Provide manual and adaptive controlling, provide stylus, buttons, gesture and voice based controlling when possible.*

   This heuristic finds some problems with Facebook - there are no additional ways to interact with the service. Users have to manually control what they are doing.

7. *Respect user's privacy and security. The world is getting more technologically oriented (e.g. electronic bank services, electronic communication services, electronic customer records). Even thought our attitudes have changed to be more open for technological solutions, we are still concerned about our privacy and security issues. User experience is always dependent of the uncertainty of how reliable the service is in terms of privacy and security.*

   User's privacy and security are respected in Facebook. However the privacy aspect has been let for the users to decide. As mentioned before in heuristic number three, the privacy settings are messy to say the least. Users have no way to be completely sure about to whom their messages and pictures are available

to. Even if the users manages to get all the privacy options correct and the messages and pictures are only visible to the intended persons, how can they be sure that these pictures are not shared forward or if someone has access to them through a mutual friend.

There is also the privacy issue with applications. Applications may demand rights to access user's information, messages and even post for the user. Applications also often wish to share user's achievements with his or her friends. Users may have problems discerning what information is shared and how it is used by the application. Users also may wonder whether the application is posting any unintended posts into the user's wall. The same fears and concerns apply for services that use Facebook identification.

This aspect is perhaps where Facebook fails the worst in these heuristics. Uncertainty in privacy settings is present in Facebook messaging. The commonly proposed solution to this is to estimate whether the message being put in Facebook is suitable to be presented publicly. This does not however remove the problem that the privacy issue is still very much at the core of Facebook design.

8. *Support user's activities but do not force them. All services that you provide to the user should be shown from a supportive perspective, e.g. how does this service support me in my actions or my everyday life. The service or product is not allowed to force a user in a one way or another. Forcing will have a negative impact on user experience.*

Facebook has a plethora of activities created for the users and it is very hard to tell how any of them, with the possible exception of messaging, support users in their everyday lives. Messaging and socializing is one of the main human needs and Facebook creates a convincing illusion that the user is actually connected to other human beings through the web service.

Facebook does not force users to do anything; in fact, Facebook only offers suggestions about things that it believes might be in the user's interest; new friends, pages that the user might be interested in and information about what games the user's friends are playing. Birthdays are hard to remember, but Facebook reminds the users on their friends' birthdays and offers a simple way to post at the walls of the friends having birthdays. Thus it can be considered that Facebook is promoting social interaction to the users.

9. *Go for a perfect visual design. From a user experience point of view, visual aspects have two meanings. The first, and most important, is that the visual design can improve usability by making the user interface more understandable, consistent and guiding. The other meaning is to make the user interface aesthetically pleasurable by designing visual aspects. It is easier to select visual elements that do not irritate the majority of the users, than to try to design solutions that are pleasant for all. In addition to these, selections in visual design, for instance use of color, can have an impact on user experience by the values one respect (such as health, fitness, nature, beauty).*

Visual design in Facebook has changed a lot in the time it's been available for users. Every time changes are made, the people are complaining how the felt experience changes for the worse. Currently Facebook's design is not perfect by a longshot; there are too many pieces of information, advertisements and visual

elements, crammed into one page. The main function of Facebook, if there are any that can be considered "main" is lost in the overload of information.

The grouping of elements however makes a small difference for the usability. Users know what these different elements stand for with perhaps the exception of the "recent friend activity" element. The advertisements no matter how easy to remove are an inconvenience that the users have to endure if they do not possess the skills to remove them from the Facebook with a browser add-on.

10. *Give a surprise gift. This means that we want more. Usability is not enough. "Jackknife mobile phone" is not enough. Breadth of experience is not allowed to decrease. Give some extra for the user. Design something extra, which makes the user happy, surpass the user's expectation, increase or improve user's experiences. User experience is the seventh sense that people want to use for sensing technology – sensing life within technology.*

This may very well be what the driving force behind the success of Facebook is. Even though there have been similar services available online for quite some time, Facebook is the only service that has allowed users to create content for it, even make money out of it. This surprise gift has been the reason behind thousands of applications being developed for Facebook alone. The availability of scores of applications has a direct impact on Facebook user amounts.

The amount of users is important for many reasons, with so many users Facebook succeeds in being the one place where its users can find all of their friends from - this in turn tempts even more people join in. With so many users, the Facebook feels more alive and active. With a vast user base, the amount of new applications designed for it also grows larger which means more to do for the masses of users.

Facebook is also updated regularly, adding more features to the already present ones. One of the more recent additions was a so called friend page, which shows the user his mutual messages and activities with a selected friend. This is one of the surprise gifts that Facebook offers for its users - the function is not in any way important or utilitarian, but it is a piece of information that may occasionally be nice to know.

3.3.2 Summarized findings (2013)

In this section the flaws gathered through heuristic evaluation are presented and summarized in tables for easier analysis and quantification. The findings will be accompanied with severity rankings (Table 3) ranging from 0 – 4, similar to those of usability evaluation. The severity rankings are adapted by the author from usability severity ranking table which is created from Nielsen's (1995) article *Severity rankings for usability problems.*

The severity rankings were adapted by the author to express another issue with user experience evaluations. Can such rankings be assigned to deviations found via UX heuristic evaluation? Because of the subjective nature of user experience (Väänänen-Vainio-Mattila & Wäljäs, 2009) the severity of found deviation is more a matter of an individual opinion than fact. The matter is discussed in more detail in the discussion Chapter of this thesis.

Table 3.   Rating scale severity rankings for user experience problems

Severity of findings	
Rating	**Explanation**
0	Violates a heuristic but is not a user experience problem
1	Violates a heuristic and is a superficial user experience problem. The problem causes minimal fluctuations in conceived experience.
2	Violates a heuristic and is a slight user experience problem. Problem causes fluctuations in conceived experience but does not affect the overall experience greatly.
3	Violates a heuristic and is a major user experience problem. The problem cases severe fluctuations in conceived experience and will lower the overall experience drawn from the service.
4	Violates a heuristic and will impose a negative user experience to anyone trying to use the service or a program.

The heuristic evaluation for Facebook's user experience also yielded thirteen deviations from the heuristics by Arhippainen (2009). There were five more severe deviations, which could impact the user experience of Facebook. These deviations were; *hard-to-understand privacy settings, uncertainty of media privacy, uncertainty of application security, uncertainty of application privacy* and *uncertainty for visibility of application's notifications* (Table 4). The uncertainty of privacy is a problem with Facebook and there are many things that might go wrong with it, these things are considered more thoroughly in the exploration section.

When compared to the results of the usability evaluation, there are three more severe problems when it comes to user experience. However, four of the five deviations categorize under the same broad heuristic - *respect users privacy and security*. So in a way, there are only two heuristics compromised, likewise in usability, when it comes to user experience of Facebook.

In the grand scheme of things, the same two heuristics which had the most severe deviations also were involved with the most deviations - counting for nine out of the thirteen found deviations. Of the three heuristics that also yielded deviations only one deviation was rated two. It is clear that there are still some issues with Facebook's user experience, but that does not change the fact that the service still remains one of the most popular services available.

One of the deviations found revealed an interesting idea, and though the deviation itself was rated zero, the notion could yield interesting results if taken care of. The required Internet access is a deviation that may not occur to many people since the service is online. And while the Internet access is eventually necessary to update the Facebook, there could be a way to work offline and update Facebook once an Internet connection is achieved.

Table 4. Summary of user experience problems

#	Problem	Severity ranking	Heuristic number	Broad Heuristic
1	Confusing for beginners	2	3	Surpass expectations
2	No tutorial available	1	3	Surpass expectations
3	Hard-to-understand privacy settings	3	3	Surpass expectations
4	The importance of friend grouping is unclear at first	2	3	Surpass expectations
5	The privacy settings need to be individually changed for each media upload	1	3	Surpass expectations
6	Internet access is required	0	5	Design the product or service to fit the intended contexts
7	Only manual interaction through keyboard and cursor	2	6	Provide several ways to interact with the product
8	Uncertainty of media privacy	3	7	Respect users privacy and security
9	Uncertainty of application security	3	7	Respect users privacy and security
10	Uncertainty of application privacy	3	7	Respect users privacy and security
11	Uncertainty for visibility of application's notifications	3	7	Respect users privacy and security
12	Too much information in the main page	1	8	Go for perfect visual design
13	Distracting advertisements	1	8	Go for perfect visual design

The two last deviations considered the visual design of Facebook. This is an interesting heuristic because the results might differ with more evaluators. Most of people have their own idea about what is visually pleasing and with a service as massive as Facebook, it is almost impossible to create a visual look that would be pleasing to everyone. This being said, there are of course things that most of people will find unappealing or even disturbing. Such as; excess amount of information and disturbing advertisements.

3.4 Heuristic evaluations with four evaluators 2009

There have been no previously published heuristic evaluations for Facebook's user experience using the heuristics by Arhippainen (2009). However, the heuristics have been used as a guideline through which Facebook was evaluated against in a course *interaction design* by Oulu University, department of information processing science. The evaluation was done by a group of four information processing science students in the year 2009. All of the students had at least a year of experience with using Facebook and at least two years of studies in the field of information processing sciences.

The results of these evaluations were looked into in this thesis due to the fact that they revealed interesting insights from the year 2009. The results were also chosen because the documentation from the evaluations is comprehensive enough so that the results could also be documented in this thesis without the possibility of data corruption. The

final reason behind the selection of the results from the evaluations from 2009 is that the main goal of this thesis is not in itself to evaluate Facebook, but instead evaluate the evaluation method used both in the year 2009 and 2013.

In this section the findings from that evaluation are presented, summarized and then compared against the findings presented in this thesis. The heuristics are presented in their broad form without the additional explanations.

3.4.1 Exploration of flaws through heuristics (2009)

The usability issues explored against heuristics by Nielsen and Molich (1994).

1. *Visibility of system status*

    Status changes generate an instant message to Facebook and e-mail if wanted.

2. *Match between system and the real world*

    There are many terms that may appear unfamiliar

3. *User control and freedom*

    There is an option for deleting

4. *Consistency and standards*

    Consistency and hierarchy is in order

5. *Error prevention*

    Changes happen immediately. Confirming to an action would be a good addition

6. *Recognition rather than recall*

    Using Facebook is consistent

7. *Flexibility and efficiency of use*

    Application boxes can be added to personal profiles

8. *Aesthetic and minimalist design*

    Facebook aims for simplistic design but is still somewhat incoherent

9. *Help users recognize, diagnose and recover from errors*

    Facebook informs the user about problems and asks to try again

10. *Help and documentation*

    No thorough help available. For example, hovering the mouse over a button could explain the function.

The user experience issues explored against the heuristics by Arhippainen (2009).

1. *Ensure Usability*

-

2. *Provide utility matching with the user's values*

   There is value for messaging and entertainment usage

3. *Surpass the user's expectations and minimize the gap between negative expectations and real usage*

   - not applicable

4. *Respect the user. Know the target user groups*

   Usability and functionality of Facebook is improved based on the feedback of the users.

5. *Design the product or service to fit the intended contexts*

   -

6. *Provide several ways to interact and leave choice for the user*

   - not applicable

7. *Respect user's privacy and security*

   Users can define what they want to share with other users

8. *Support user's activities but do not force them*

   - not applicable

9. *Go for a perfect visual design*

   Simplistic design

10. *Give a surprise gift*

    - not applicable

3.4.2 Summarized findings (2009)

In this section, both usability and user experience findings from 2009 by Huhtala, Isotalus, Itkonen and Laitinen are presented. The findings are tabled and accompanied with severity rankings similarly to the previous summaries.

Table 5. Summary of usability problems

#	Problem	Severity ranking	Heuristic number	Broad Heuristic
1	Unfamiliar terminology	1	2	Match between system and the real world
2	Immediate changes	1	5	Error prevention
3	No confirmation for actions	3	5	Error prevention
4	No thorough guidance	1	10	Help and documentation

In 2009 the amount of found usability issues was low - only four deviations were discovered at the time (Table 5). There are certain reasons behind this; the expertise of the evaluators, the nature of the evaluation process and the heuristics being too general for evaluating the type of service that is Facebook.

But as there are reasons for the number of deviations being as low as they are, there are several reasons why there could and ought to be more. Firstly, the study was done almost five years ago, in the time Facebook has been around it is certain that many heuristic deviations have been discovered and taken care of. However five years ago some of these deviations should have still been uncovered, unless the service itself has not been improved for the last five years. Secondly not all the evaluators were familiar with the service. This could potentially both increase and decrease the found deviations to some degree. For a beginner to evaluate such a massive service, it is hard to understand what is intentional and what to consider a deviation from the heuristics.

All of the above said, it is hard to argue against the fact that the expertise of the evaluators and the evaluation situation is the key factor contributing to the small number of findings. There is one deviation which did not remain unnoticed even from the inexperienced evaluators –*"no confirmation for actions"*. This in an intriguing deviation because it has been around for at least five years and still has not been fixed. And yet it is rated amongst one of the highest of all found deviations.

Table 6. Summary of user experience problems

#	Problem	Severity ranking	Heuristic number	Broad Heuristic
1	Unfamiliar terminology	1	9	Go for perfect visual design

User experience issues seemed to be even scarcer in 2009. The only deviation which was found was the *unfamiliar terminology* (Table 6). Even though this is a found flaw in the service, it is hardly any reason for concern because it can be overlooked and is hardly more than a minor nuisance. The lack of found deviations can be credited to the same reasons than in usability; inexperienced evaluators and too general heuristics to evaluate the type of service.

3.5 Combined and compared results

This section will consist of a table (Table 7) summarizing the findings from all the heuristic evaluations and a short analysis of the immediate findings that can be drawn from the findings.

Table 7. Found deviations from heuristics

#	Heuristic type	Problems	Highest severity rating	Problem
1	Usability	13	3	No undo or redo. No confirmation required for the unfriend function.
2	User Experience	13	3	Hard-to-understand privacy settings. Uncertainty of media privacy. Uncertainty of application security. Uncertainty of application privacy. Uncertainty for visibility of application's notifications.
3	Usability (2009)	4	3	No confirmation for actions
4	User Experience (2009)	1	1	Simplistic design

The summarizing table (Table 7) points out one crucial aspect about heuristic evaluations – experience of the evaluators. As Doubleday et al. (1997) and Ling and Salvendy. (2005) have pointed out, the reliability and thoroughness of heuristic evaluation is directly proportional to the expertise of the evaluators.

In year 2009, despite the fact that the amount of evaluators was higher than in the recent analysis, the number of found faults was considerably lower. The most severe issues of user experience evaded the evaluators' sight. However this could also mean that the user interface and the service itself have gone through some changes which have taken the service further from that of the 2009 user experience considered.

It may be impossible to compare these researches and that is not the main focus of this thesis. This comparison shows however that in order for the heuristic evaluation to work, the evaluators need to have advanced knowledge of usability or user experience – whichever is under evaluation. The effects of the evaluators' knowledge is better seen in the heuristic evaluation of user experience, where four of the ten heuristics are left untouched, marked as non-applicable. There are certain difficulties in applying the heuristics to a service that has already been in use for many years and has reached a huge mass of users.

4.   Reflection

This section will assess the parallels between the literacy findings and the conclusions drawn from the empirical results of the heuristic evaluations. The results of both usability and user experience evaluations are taken into consideration to find the advantages and shortcomings of heuristic evaluation for user experience.

4.1 Methodological challenges

The literature lists at least six challenges associated with heuristic evaluations; the need for couple of evaluators (Jeffries & Desurvire, 1992; Nielsen, 1994; Hertzum & Jacobsen, 2001), reliance on the sole expertise of the evaluators (Doubleday et al., 1997; Jeffries & Desurvire, 1992; Kantner & Rosenbaum, 1997; Ling & Salvendy, 2005), false-positive errors (Cockton et al., 2008; Jeffries & Desurvire, 1992), selection of the right set of heuristics (Holzinger, 2005), separation from the actual end users (Kantner & Rosenbaum, 1997; Holzinger, 2005) and the inability to offer any support for countering the deviations (Ling & Salvendy, 2005). All of the listed challenges were visible in the case studies to some degree.

4.1.1 Number of evaluators

There is a need for multiple evaluators for a couple of reasons. Five evaluators have been estimated to find around two thirds of all usability issues with the best cost-to-benefit ratio (Nielsen, 1994a). In the case studies 13 deviations were found from both usability and user experience evaluations. Nielsen and Molich (1990) have estimated that a single evaluator can find between 20 – 51% of the problems but depending on the expertise of the evaluator and experience with similar systems, the number could be even higher (Figure 3).
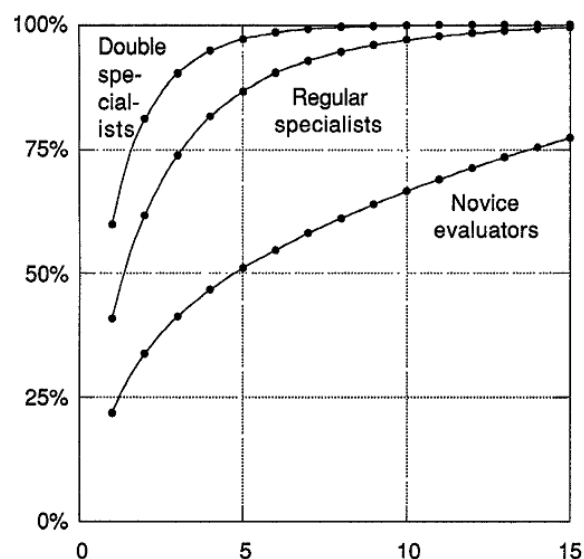

Figure 3. Proportions of usability issues found in function to the number of evaluators (Nielsen, 1992)

It is hard to imagine that the found thirteen deviations would only count for around 23% of total faults. It still must be assumed that with only one evaluator and such a short evaluation timespan, many of the existing faults still remain undisclosed. Hertzum and Jacobsen (2001) state that overlap between two evaluators estimates around $5 - 65\%$, which means that the possible deviations found with two evaluators would estimate around $13 - 21$. The possible increase of 65% is too significant to ignore which means that heuristic evaluation should always include at least two evaluators.

There is also another aspect supporting the need for multiple evaluators. As of now, the severity ratings assigned to found deviations are only proposals. With several evaluators the arithmetic mean between all proposed severity ratings could be counted and thus it would make the assignment of severity values more reliable.

4.1.2 Expertise of evaluators

One of the main discoveries between the evaluations from 2009 and this thesis was the impact of expertise of the evaluators. In the year 2009 only four deviations were found in the usability evaluation and just one deviation from the user experience evaluation. This could indicate two possibilities; either the quality of the service has decreased drastically in both usability and user experience. Or the expertise of the evaluators was not as high in 2009 as it is around the time of writing this thesis.

The level of expertise is most clearly visible in the user experience evaluation from 2009 in which four out of the ten heuristics are left inapplicable and two are left blank. Out of ten heuristics, only four were applied. Doubleday et al. (2005) have stated that less experienced evaluators will not find as many deviations or may not consider the found deviations to be faults. This is also visible in figure 3 where the expertise of the evaluators is also shown in the graph. The final impact of the expertise is the differences in severity ratings – whereas more experienced evaluators may be more agreeable on the severity of a found deviation, between the less experienced evaluators, the results may vary profoundly (Doubleday et al., 2005).

4.1.3 False-positive errors

Heuristic evaluation has been credited with finding the most severe problems with the least effort. (Jeffries et al., 1991). However it is also found to discover many false-positive errors (Cockton et al., 2008; Jeffries & Desurvire, 1992). This notion is hard to verify in the case study without conducting thorough usability and user experience studies about Facebook with actual users and comparing the list of found deviations with the results of those studies.

There are some deviations that can be considered false-positive even though they fulfill the properties of heuristic evaluation faults. For example, the heuristic number three; *no titles on all of the pages* is an example of a deviations which can be attributed as such but a normal sane user may not see it as a deviation. Normal user knows where he or she has navigated in the service and thus does not need to be reminded where he or she is. However it is still a deviation from the heuristics because users should be visible of system status and knowing whether they navigated into a part of the service they meant to.

Some problems can be considered false-positive in a sense that they do not show up in user testing. Minor problems such as the three steps required to delete an unwanted post may not be revealed in user testing. But whether this problem is a false-positive error is

questionable. Inconsistencies and inflexibility issues may reduce the efficacy of the system and cause minor fluctuations in user experience and thus should be considered a true deviation. (Nielsen, 1992.)

4.1.4 Selection of heuristics

Nielsen and Molich (1990) wanted to ease the evaluation of user interfaces and thus created the heuristics that would be simple to use and easy to apply in almost every situation. For a time they were successful and heuristics became one of the worlds most applied method of assessing usability (Rosenbaum et al., 2000; Holzinger, 2005; Ling & Salvendy, 2005). This was all in the nineties – the systems and services created at the time were quite similar to one another with the occasional exception. This meant that the heuristics were general enough to encompass almost all the systems at the time.

Times have changed and emergence of hedonic systems changed the scene. Systems with hedonic attributes were hard to evaluate using the original heuristics. User experience also changed the landscape and while Nielsens (1994a) heuristics were still one of the most used methods to assess usability, there were systems and services that simply could not be evaluated using the original heuristics.

New heuristics were created, heuristic evaluations for assessing playability (Desurvire, Caplan, & Toth, 2004), heuristics for evaluating user experience for web services (Väänänen-Vainio-Mattila & Wäljas, 2009), heuristics evaluating general user experiences (Arhippainen, 2009) and many others. As Holzinger (2005) has stated, the selection of the right set of heuristics is critical for the success of the heuristic evaluation. With the ever increasing number of different heuristics for different types of systems we are taking a step backwards into the situation of the 90's - where there were thousands of guidelines to evaluate usability of an interface.

In this case study it was found that neither set of the used heuristics was optimal for evaluating Facebook. This was also shown in the heuristic evaluations of 2009 where only some heuristics were able to be applied. In this thesis, all the heuristics were applied to a degree which they could. The meanings of many heuristics in the context of the evaluated service were carefully considered and if they could be applied, they were. If Facebook was to be evaluated again, it is worthwhile to assess which heuristics would be best suited for its evaluation.

4.1.5 End-user separation

Even if the evaluators may find out problems that would not occur to the primary users, they are not always the actual users of the system or service. This means that the evaluators may be oblivious to some things the main users find distressing. The primary users often find surprising things that would have never occurred to the evaluators. (Kantner & Rosenbaum, 1997)

In exploring the third user experience heuristic, the messaging system of Facebook is given credit. The messaging is said to encompass the needs of the users; this is in a sense - a lie. Because there were no actual users involved in evaluating Facebook, there are no guarantees that the messaging system actually deserves credit. It is possible, yet very unlikely, that majority of the users want to be able to talk to complete strangers via the Facebook chat, and the absence of that functionality is a huge usability and user experience issue.

This problem is evidently a challenge for heuristic evaluation. Some of the problems found may even be things that make the service more attractive. But even when the end-user separation can cause some issues when evaluating the usability or user experience of a system, it can also be one its strengths. The separation from end users gives more freedom to conduct usability and user experience evaluations for the evaluators. The end-user separation is what makes heuristic evaluation fast, easy and cheap. (Holzinger, 2005)

4.1.6 Deviation countering

Heuristic evaluation does not include any proposals to fix the found deviations (Ling & Salvendy, 2005). This is when the expertise and innovative abilities of the evaluators become really important. When evaluating Facebook, many of the issues found seem easy to correct, keyboard shortcuts to support undo (ctrl – z) and redo (ctrl – y) should be rather easy to implement and would remove one the three – rated deviations.

It is quite impossible to add any kind of proposals to fix problems that are discovered using heuristic evaluations. This is due to the fact that the evaluation process may uncover issues that are harder to counter than for example *no undo or redo*. This is not a huge challenge and it is not an issue that should be looked into or tried to be fixed. The main idea behind heuristic evaluations is to find deviations in the usability or user experience of the system, then make these deviations known to the people with the needed abilities to fix the found issues.

Deviation countering, however, is an issue that needs to be known. It is also an issue that the possible third party inquiring for a heuristic evaluation needs to be aware of. Heuristic evaluation is not an all resolving solution; it is merely a tool for finding things that affect the usability or user experience of a system.

4.2 Methodological advantages

Despite the shortcomings involved in using the heuristic evaluation method, there are also advantages in using the method. These advantages include aspects such as; cheap cost, relatively fast completion and easy utilization (Nielsen & Molich, 1990; Iannella, 2009).

Other benefits include the possibility of heuristic evaluation being thorough. This however requires preferably three to five evaluators when conducting usability evaluation (Nielsen, 1994a). Heuristic evaluation has also been found to find the most serious problems with the least amount of effort (Jeffries et al., 1991).

The last benefit of heuristic evaluation for systems usability is that it can be conducted to paper prototypes or other preliminary models of the designed system – long before the actual system itself is developed. (Nielsen, 1992.)

4.2.1 Cheap costs

Empirical usability studies are usually costly (Jeffries & Desurvire, 1992). They require many users, compensation for those users and a place where to conduct these studies. Nielsen and Landauer (1993) have stated that five test users are sufficient to find estimated 75% of all usability issues. There has been some debate about whether the Landauer and Nielsen's formula holds true to all situations (Woolrych & Cockton,

2001) and in a study by Spool and Schroeder (2001) the five users only managed to find 35% of all problems. This can be attributed to different types of software emerging over almost a decade since Nielsen and Landauer (1993) published their formula.

However, if five users are not enough – how many does it take? The exact amount is not an easy thing to come up with. Woolrych and Cockton (2001) state that number of needed test users depends on the users' ability to find errors, the complexity of user tasks and the tested tool itself. To sum it up, five users may or may not be enough. To ensure that the results of user testing are thorough enough it may even take 23 test users to find 86% of all problems. (Woolrych and Cockton, 2001.)

However, usability testing is expensive. Mantei and Teorey (1988) have estimated that the

"costs required to add human factors elements to the development of software"

can be as high as $128,330. Even though this sum might have been changed over the decades since it was estimated it still gives a good idea about the expenditures that usability studies may require. Nielsen (1994b) has reduced the total sum presented by Mantey and Teorey (1988) by $11,520 by reducing the number of test users to three instead of five. Woolrych and Cockton (2001) state that additional users increase costs with the additional recruitment costs, rewards, test sessions, data analysis and report synthesis. Even if most of these costs are static, data analysis becomes more complex with every new user; from this we can deduce that the costs of adding new users may be non-linear and increase rabidly with every new user.

In heuristic evaluation, the costs derive mostly from the expenditures related to the recruitment of experts to do the evaluation(s). Nielsen (1994b) has estimated the cost of bringing a usability expert to do a heuristic evaluation to be around $3,000. The sum will of course depend on the expert and the complexity of the system and therefore the time needed to do the evaluation. If we keep the $3,000 as the price tag and bring five evaluators to do a heuristic evaluation, the cost would be $15,000. In comparison, bringing five test users and keeping the cost of their recruitment linear, it would total up to $19,200 – keeping in mind that the actual cost is more likely to be higher and five users may or may not be enough. Heuristic evaluation is therefore cheaper than standard user testing. For cost-benefit ratio Nielsen (2001) gives an example of project where the cost-benefit ratio was ~48; the cost of the evaluation $10,500 and the benefits around $500,000.

4.2.2 Fast utilization

Heuristic evaluation is a rabid process (Nielsen & Molich, 1990). This is best shown with an example; the heuristic evaluations done for this thesis took two weeks to complete – this includes the exploration of faults and compiling the summarizing tables. The additional explanations took one more week but in a real evaluation situation those would not have been necessary. Thus the time required to do the heuristic evaluations for this study was two weeks.

Had there been more than one evaluator, the time would have been longer but not significantly. The reason for slightly longer completion time is that the evaluators would've had to come to a common understanding of the found deviations and their respective severities. Since heuristic evaluations are done individually, all the evaluators would have been able to do their individual evaluations alone it can be assumed that each evaluator would need from two to three weeks, which includes reporting and fixing

suggestions depending on their expertise and thoroughness (Kantner & Rosenbaum, 1997). After all of the evaluators had finished their evaluations, the results would be aggregated which would most likely take another week. This would take the amount of time needed to do two heuristic evaluations to a total of four to five weeks. The times are summarized in the table

If the evaluation would have been done in a traditional empirical user testing method, the timetable would have been a lot longer. First there would have been the creation of tests, including the tasks that need to be completed during the tests – which is critical because generally the users need to complete real tasks (Nielsen, 1996; Battleson, Booth & Weintrop, 2001). Second would have been the need to choose the criteria through which the test-users would have been chosen – then would have started the recruitment process to find suitable testers – which is time consuming even with the services currently available (Kantner & Rosenbaum, 1997).

After a suitable sized pool of candidates had been comprised, the selection process would ensue; the candidates would be evaluated against the selection criterion. Then the selected candidates would have been contacted and times for the test sessions would have been chosen. The best case scenario is that all of the tests could be done in a single day. In the worst case scenario, the tests could take as many days as there are users – keeping in mind that five users may or may not be enough. After the tests had been conducted, the results would need to be analyzed and the findings qualified and quantified. The time needed to do this depends on the number of test-users and the amount and variety of problems they discovered – for this step, usability and / or user experience experts are required. The last tasks to be done would be collecting and aggregating the findings and suggesting possible fixes and then reporting.

The literature concerning the times it takes to complete the different steps of the user testing is minimal and therefore the estimates have been used in this thesis. Though the times are mostly rough estimates, they have been tried to be kept minimum. An example case study for usability testing took 13 weeks to complete (Chisman, Diller & Walbridge, 1999). In the case study conducted by Jeffries et al. (1991) the usability evaluation done by heuristic evaluation took a total of 33.5 hours, roughly a week, depending on the daily workload put into the evaluation process. The same evaluation done by user testing took 125 hours, with twenty end-users.

From these two calculations it can be estimated that even with the lowest possible times for a standard usability testing and the highest possible time for the heuristic evaluation, the time required for the latter is greatly less. The time required to do the heuristic evaluations for this study was two weeks - thus it can be said that heuristic evaluation is much faster than standard usability evaluation with test users.

4.2.3 Easy utilization

What makes heuristic evaluation easy? When Nielsen and Molich (1990) created the heuristics their idea was to create rules of thumbs that can be used to evaluate an interface against. The general idea was that even a novice should be able to follow the heuristics and evaluate any given interface against them.

Looking at the steps required for completing a heuristic evaluation and a standard usability evaluation the difference is obvious. Single evaluator can complete heuristic evaluation in two steps; do the evaluations – basically follow a list of principles and mark the deviations. Then aggregate the results, provide fixes if required and report

them. Standard usability evaluation requires a lot more effort; multiple additional tasks such as the arduous user recruiting (Battleson et al., 2001).

One great advantage that heuristic evaluation has against many other ways to evaluate usability or user experience is its indifference towards the physical environment. It does not matter where the evaluation is conducted; evaluations conducted home are just as valid as evaluations conducted at laboratory settings or at an office. Of course there are systems that can only be evaluated in certain locations. An example for such a system could be an internal system for an enterprise that can only be accessed through the company Ethernet.

However, for a system that can be accessed from anywhere, it doesn't matter where the actual evaluations take place. This means that evaluators don't need to worry setting up laboratory surroundings for the evaluations. Of course when the evaluators have no experience or knowledge about the evaluated system, it is worthwhile to have a domain expert observer watching and helping the evaluator if there is a problem. (Nielsen, 2001.)

Case study that was conducted for this thesis also supports the notion of heuristic evaluation being easy to conduct. Even for an undergraduate student evaluating web services usability and user interface, through heuristics that were not exclusively designed for such evaluations, the process was relatively easy. Some heuristics provided with a challenge in applying them but in general the process was non-abrasive.

4.2.4 Thoroughness

Heuristic evaluation can be thorough; even three evaluators are sufficient to find around 60% of total deviations (Nielsen, 1994a). This aspect was unable to be confirmed in the case study done for this thesis because there was only one evaluator. However, the found 13 deviations itself supports the notion of heuristic evaluation being able to quickly harvest the so called "*low hanging fruits*" (Kantner & Rosenbaum, 1997).

The thoroughness of Heuristic evaluation is also highly dependent on the expertise of the evaluators. In the evaluations conducted in 2009 with four evaluators, only total of five deviations were found in both usability and user experience evaluations. This supports Nielsens (1992) estimates of deviations found per evaluator. Four evaluators have been estimated to find slightly more issues than a single regular expert, but being estimates, this number is still due to debate.

4.2.5 Most serious problems with least amount of effort

Heuristic evaluation is said to find most serious problems with the least amount of effort (Jeffries et al., 1991). Desurvire et al. (1992) support this theory when evaluators are considered experts. It means that heuristic evaluation reveals problems with the highest severity ratings. This is due to the fact that while doing the heuristic evaluation, evaluators often tumble into the most severe problems because of their utilization inhibiting - and frustrating nature.

The case study within this thesis supports the notion of most serious problems found through heuristic evaluation. The highest found problems were rated three in the severity scale – problems that cause frustration and can cause serious issues while using the service. It was to be expected that four's from the severity ranking would not come

up during the heuristic evaluations because of their utilization crippling nature – such problems should have been fixed in a decade in such a popular service.

There is no certainty that the highest ranking deviations were mostly found but the ones that were found, did so without any inconveniences. More research however is warranted to compare heuristic evaluations against other methods of evaluation to further study the statement.

4.2.6 Early in the design process

The final advantage of heuristic evaluation covered in this thesis is its ability to be able to be used even before the actual development of a system or service (Nielsen, 1992). The ability to discover problems before the development of the system makes design process a lot cheaper because problems are more costly to fix the further the development is (Boehm, 1988).

The applicability in early phases does however suit better to the usability than user experience evaluations. This is simply due to the different nature of the two. Usability is more about finding flaws regarding technical aspects of the system (Bevan, 2009) whereas user experience is more concerned about the wider relationship of the user and the system (McNamara & Kirakowski, 2006; Law et al., 2009). Thus the results may not be as rich as in the usability heuristic evaluation – this does not mean that user experience heuristic evaluations in the early design phase are any less valid. The deviations affecting systems user experience found in the early UX heuristic evaluations can prove highly advantageous in the later design phases.

5.   Results and discussion

In this Chapter this thesis will present the results and discussion of the possible problems about the advantages and challenges of heuristic evaluation of systems usability and the heuristic evaluation of systems user experience. With the focus slightly in favor of the user experience heuristic evaluation. There are some concerns that need to be taken into consideration before stating that the same advantages and challenges apply to both evaluations.

5.1 Challenges of UX heuristics

It is well known that the "*silver bullet*" heuristic evaluation had many challenges; the need for three to five evaluators (Jeffries & Desurvire, 1992; Nielsen, 1994; Hertzum & Jacobsen, 2001), reliance on the sole expertise of the evaluators (Doubleday et al., 1997; Jeffries & Desurvire, 1992); Kantner & Rosenbaum, 1997; Ling & Salvendy, 2005), abundance of false-positive errors (Cockton et al., 2008; Jeffries & Desurvire, 1992), choosing appropriate heuristics (Holzinger, 2005), end-user separation (Kantner & Rosenbaum, 1997; Holzinger, 2005) no fix suggestions for found deviations (Ling & Salvendy, 2005). All of the above and some additional ones also challenge the utilization of the heuristic evaluation for user experience.

Väänänen-Vainio-Mattila and Wäljas (2009) have shown in their article that it is hard to create user experience heuristics that can cope with all the different varieties of web services. They had created seven heuristics to evaluate user experience of web services of three different types. The creation process has been greatly influenced by these three services. Two of the seven heuristics were only applicable to one particular service.

A challenge with regular usability heuristic evaluation was that the experts were not the actual users of the system, and it was inevitable that the real users managed to find errors that remained unnoticed by the experts (Kantner & Rosenbaum, 1997). User experience being even more of a subjective issue, it might be assumed that the experts are no better at finding user experience issues than they are at finding usability issues. Based on the findings of Kantner and Rosenbaum (1997) other methods are also required to find the issues that will cause dissatisfaction to the actual end users.

Other finding by Väänänen-Vainio-Mattila and Wäljas (2009) was that the experts asked to evaluate the services through the heuristics found it hard to evaluate user experience for other people. Evaluating usability issues is a much more straightforward issue than evaluating user experience. This is because user experience is a wider, individual relationship between the user and a product or service (McNamara & Kirakowski, 2006; Law et al., 2009).

The subjectivity raises in intriguing notion about the number of evaluators required to conduct a heuristic evaluation. Nielsen (1994a) has advocated that three to five evaluators are able to find two thirds of all usability problems and assign them with fairly reliable severity ratings. How many evaluators are needed to find user experience issues that cause dissatisfaction, and how many evaluators are needed for reliable consensus?

Heuristics also involve another dilemma, Vermeeren et al. (2010) propose that user experience evaluations should be done beforehand and during interaction with the object or service, as well as after the interaction has ended. However, heuristic evaluation during the use of the service can hinder the feeling of flow – explained in more detail in (Csikszentmihalyi, 1988), which is often seen as an important factor (Hassenzahl, 2008, Deng, Turner, Gehling, & Prince, 2010) when determining user experience.

Another drawback about heuristic evaluation is that it was originally created to evaluate usability which remains constant before, meanwhile and after using the product or a service. One of its greatest merits lies in evaluating early (Nielsen, 1992) and correcting the faults before the development of the system – thus reducing the costs associated with bug fixes later in the development phase (Boehm, 1988). While this can be achieved with usability evaluations, user experience changes over time (Mendoza & Novick, 2005; Karapanos et al., 2009) and what may have once felt like a novel and an interesting way to accomplish things may eventually frustrate users and thus reduce the changes for a prolonged usage. The first (*ensure usability*) and the fifth (*design the product or service to fit the intended contexts*) UX heuristic (Arhippainen, 2009) reflect the idea that the service or a system is intended to be used after the initial state of incorporation but are they enough to ease the transition between all the three stages? (Karapanos et al., 2009.)

Thus heuristic evaluation does not provide with a clear cut vision of a systems or services user experience over time. Different aspects of the system may eventually need to be changed even if the initial heuristic evaluation shows no issues with the systems user experience. Väänänen-Vainio-Mattila and Wäljas (2009) have suggested evaluating systems user experience in different stages of utilization to get a better image of the temporality of user experience, however they also agree that it would be a trade-off with heuristic evaluations being fast and a low-cost evaluation method.

In their research "*Towards the evaluation of UX*" Ardito, Costabile, Lanzilotti & Montinaro (2007) have studied the evaluation of usability and user experience in games. They have come to the conclusion that there are multiple methods available to assess both dimensions but there is no one method that can accomplish everything. However they also argue that heuristic evaluation is not applicable to games because expert evaluation would not produce as interesting notions about the game.

This thesis raises a question whether such a general heuristics be created that they encompass all the possible services, whether they are web services or conventional products – and still maintain the viability and applicability to cover a single service or a product. Additional questions within the topic might consist of such questions as, how can we measure experience for other people when user experience is, as told before, an individual relationship between the user and product or service.

The first UX heuristic (Arhippainen, 2009) "*Ensure usability*" is very broad. Even though this is not as much a problem, it is good to take into consideration the fact that the first heuristic could possibly require a usability evaluation just to make sure that it is fulfilled. In this thesis, the first heuristic was applied by evaluating Facebook's usability through Nielsen's (1994) heuristics. The first heuristic could be considered a reminder that usability evaluation, whatever it may encompass, should be done before user experience can be evaluated – since user experience is intertwined tied to the usability of the system (Hassenzahl, 2005).

5.2 Advantages of UX heuristics

Even though UX heuristic evaluation certainly has its challenges, there are also numerous advantages in using the method. Since the method is procedurally similar to a traditional usability heuristic evaluation, the same advantages concerning the evaluation itself are static. The main advantages still remain the relatively cheap cost, the ease and speed at which heuristic evaluations can be conducted (Nielsen & Molich, 1990; Iannella, 2009).

From the case study it can be construed that heuristic evaluations, whether usability or user experience, managed to find rather severe issues – ranked three on the severity scale. It was expected that there would be no higher ranking issues because of their crippling, consumption preventing nature. The ratings can be questioned of course, there was only one evaluator assigning the severity rankings which brings us back to the need for several evaluators. This also raises another issue whether such ratings can even be assigned to user experience issues. Being very subjective in nature (Väänänen-Vainio-Mattila & Wäljas, 2009) other evaluators and end users might find issues ranked high in the severity scales much less disturbing or not even an issue at all. However high the found deviations may rank in the severity scale, they were indeed found with the least amount of effort. Just mere browsing, keeping in mind the heuristics is enough to uncover the biggest and easy-to-find issues, the so called *"low hanging fruit"*.

Heuristic evaluation can also be done to early prototypes (Nielsen, 1992) and it will therefore reduce the costs of bug fixes (Boehm, 1988). With both, usability and user experience heuristic evaluations being procedurally similar, UX heuristic evaluation is also able to be conducted to early prototypes. With early prototypes, heuristic evaluations ability to find the low hanging fruit is of most advantageous. Getting rid of the usability or user experience crippling issues before the development of the system can reduce the costs from fixing the things in the actual development process. It is still uncertain how the effects of heuristic evaluation will affect the experiences over time but heuristic evaluation is a great way to eliminate major mistakes that would cause issues with the initial user experience.

Heuristic evaluation of user experience also takes into account the unexpected needs that users may have. As Arhippainen (2009) states:

> *"User experience is the seventh sense that people want to use for sensing technology – sensing life within technology."*

Users expect that certain features will be available with the systems or services they invest into. The lack of these features will cause dissatisfaction but their existence alone does not guarantee satisfaction. (Oliver, 1997; Jokela, 2004.) UX Heuristic evaluation steers towards surpassing the expectations of the users, thus trying to provide the users satisfaction. There is also one additional positive issue regarding the UX heuristics; they can also take into consideration the unexpected needs of the users by providing them with something extra (Arhippainen, 2009, pp. 224). Users expect certain features to be available (Jokela, 2004) but for the system to provide a good user experience the extra functionality is required – whatever it may be.

The advantages that heuristic evaluation boasts are fewer in comparison with the challenges it poses – especially when evaluating user experience. However, it needs to be kept in mind that the advantages are very high status, and many of the challenges mere footnotes. Being easy, fast, cheap, and able to be conducted early in the development, heuristic evaluation is very worthy method of assessing either, or both, usability or user experience.

6.   Conclusion and future work

This thesis aimed to evaluate the advantages and challenges of using the user experience heuristic evaluation method. The aim was also to provide the scientific community with arguments for and against using the method. There were multiple challenges that were already present with the original heuristics by Nielsen and Molich (1990) and all of the previous challenges, with a few additions, were applicable to the user experience heuristics created by Arhippainen (2009). However the original advantages weigh heavily which brings the scales to equilibrium – the decision to use heuristic evaluation therefore depends on the system and other external factors. The concluded results will be explored in this Chapter.

User experience became a buzzword in the scientific community ever since its conception in the mid 90's (Hassenzahl & Tractinsky, 2006; Ardito et al., 2007). This was partly due to the fact that the traditional concepts of usability were deemed too narrow to fully explain the complex nature of HCI (Law et al., 2009). However the scientific community could not come to a general understanding of the term and what it encompassed – which lead to a plethora of different definition and assumptions about what user experience is, should be and most importantly how to evaluate for. (Law et al., 2007; Hassenzahl, 2008.)

There are many methods to assess usability; heuristic evaluation (Nielsen & Molich, 1990), cognitive walkthrough (Wharton, Rieman, Lewis & Polson, 1994), formal usability inspections (Kahn & Prail, 1994), empirical end user testing and many others. Of these methods, heuristic evaluation also received the buzzword status and the boldest of researchers have stated that traditional testing was no longer required due to the silver bullet of all evaluations – heuristic evaluation (Jeffries & Desurvire, 1992). However, it is well known that there is no silver bullet.

Heuristic evaluation still remained highly used due to its many advantages; cheap, fast and easy to conduct (Nielsen & Molich, 1990; Iannella, 2009). It was also found to find the most severe problems with the least amount of effort (Jeffries et al., 1991) and it could be conducted to paper prototypes or other preliminary models of the designed system – in the design phase before the actual development of the system (Nielsen, 1992).

Arhippainen (2009) created a set of general heuristics for a systems user experience and this thesis was conducted to explore their possible advantages and challenges. The advantages that were studied for the original heuristics are also present for the heuristic evaluation of systems user experience – however the thoroughness on which heuristic evaluations boast on could not be verified in the case study due to the simple fact that there was only one evaluator.

One challenge found was the temporality of user experience that the heuristic evaluation cannot evaluate for. User experience changes over time (Mendoza & Novick, 2005; Karapanos et al., 2009). There are three stages identified in systems usage; orientation, incorporation and identification (Karapanos et al., 2009.). It has been suggested that in order to get a better vision of systems temporal user experience, heuristic evaluation should be carried out in different stages of the systems usage – it was also noted that doing so would contradict two out of the three original advantage of heuristic

evaluation; the cheap cost and fast utilization. (Väänänen-Vainio-Mattila & Wäljas, 2009)

The final major issue is the subjective nature of user experience issues. When asked to evaluate systems user experience, the experts found it increasingly difficult to wear the end-user hat (Väänänen-Vainio-Mattila & Wäljas, 2009). This is due to the fact that some of the issues found can even be desired in the system, even though heuristics may classify them as deviations. This is partly due to the problems in the concept of usability evaluation being mostly able to only find deviations from the norm instead of accepting the possible novel ideas (Greenberg & Buxton, 2008).

With all of the above being said it is never a good idea to trust a single methodology alone; no single evaluation method can provide all the insights into a system and they are best used in conjunction with one another (Jeffries & Desurvire, 1992). Thus heuristic evaluation has its rightful place amongst the other methods such as user testing.

The restrictions of this thesis lie mostly in the case study. The evaluations should be done with more than one evaluator to get a better image of the total amount of systems deviations and a better consensus on their severity. The other restriction is the splintered nature of user experience definitions and assumptions (Law et al., 2007) due to which there are not as much literature to cover all the issues with user experience. One great restriction of this thesis is the negligible literature about the time estimations for the steps required to do user testing, therefore the times proposed in this thesis are nothing more than approximate estimations. Lastly, the UX heuristics by Arhippainen (2009) have not been studied critically in any studies which makes assessing them that much harder.

The thoroughness of user experience heuristics could not be assessed in this thesis due to the fact that there is no empirical research on the matter. The evaluation conducted in this thesis shows an approximation that it can be as thorough as usability evaluation – based on the single fact that they both managed to find an equal amount of issues. More research is badly needed – empirical studies in particular, to evaluate the thoroughness of the heuristics.

Research is also warranted into finding out how many evaluators are required to find Nielsen's (1994a) two thirds of all available deviations and also how many evaluators are required to reach consensus on the severity of these found issues due to the subjective nature of user experience. The severity scales and their applicability to user experience issues is also a worthwhile research venue. Additional research is also needed to find out whether heuristic evaluation can be adapted to support temporal effects of user experience. Additional questions regarding the issue could be whether such an adaptation is possible within a single heuristic evaluation and whether additional heuristics or even a new set of heuristics are required?

References

Ardito, C., Costabile, M., Lanzilotti, R., & Montinaro, F. (2007). Towards the evaluation of UX. In *Workshop "Towards a UX Manifesto" at the International Conference HCI 2007, Lancaster, UK* (pp. 6-9).

Arhippainen, L. (2009). Studying user experience: issues and problems of mobile services--Case ADAMOS: User experience (im) possible to catch. *Fac. of Science, Dep. of Inf. Proc, Univ. of Oulu, Finland*.

Arhippainen, L. (2013) A Tutorial of Ten User Experience Heuristics. Tutorial in Academic MindTrek Conference, 1.10.2013. ACM Press, 336-337.

Battarbee, K., & Koskinen, I. (2005). Co-experience: user experience as interaction. *CoDesign*, *1*(1), 5-18.

Battleson, B., Booth, A., & Weintrop, J. (2001). Usability testing of an academic library web site: a case study. *The Journal of Academic Librarianship*, *27*(3), 188-198.

Bevan, N. (2009). What is the difference between the purpose of usability and user experience evaluation methods. *In EXAM 09 Workshop,* INTERACT.

Bevan, N., Kirakowski, J., & Maissel, J. (1991). What is Usability?. In *Proceedings of the 4th International Conference on HCI*.

Boehm, B. W. (1988). Understanding and controlling software costs. *Journal of Parametrics*, *8*(1), 32-68.

Chisman, J., Diller, K., & Walbridge, S. (1999). Usability testing: A case study. *College & research libraries*, *60*(6), 552-569.

Cockton, G., Woolrych, A., Lavery, D., Sears, A., Jacko, J., Tsuchiya, I. & Grandy, G. (2008). Inspection based evaluations.

Csikszentmihalyi, M. (1988), "The flow experience and its significance for human psychology*", in Csikszentmihalyi, M., Optimal experience: psychological studies of flow in consciousness*, Cambridge, UK: Cambridge University Press, sivut: 15–35,

Deng, L., Turner, D. E., Gehling, R., & Prince, B. (2010). User experience, satisfaction, and continual usage intention of IT. *European Journal of Information Systems*, *19*(1), 60-75.

Desurvire, H., Caplan, M., & Toth, J. A. (2004). Using heuristics to evaluate the playability of games. In *CHI'04 extended abstracts on Human factors in computing systems* (pp. 1509-1512). ACM.

Desurvire, H., Kondziela, J., & Atwood, M. E. (1992). What is gained and lost when using methods other than empirical testing. In *Posters and short talks of the 1992 SIGCHI conference on Human factors in computing systems*(pp. 125-126). ACM.

DIS, I. 9241-210: 2010. *Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems*.

Donahue, G. M. (2001). Usability and the bottom line. *Software, IEEE*, *18*(1), 31-37.

Doubleday, A., Ryan, M., Springett, M., & Sutcliffe, A. (1997). A comparison of usability techniques for evaluating design. In *Proceedings of the 2nd conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 101-110). ACM.

Forlizzi, J., & Battarbee, K. (2004). Understanding experience in interactive systems. In *Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques* (pp. 261-268). ACM.

Greenberg, S., & Buxton, B. (2008). Usability evaluation considered harmful (some of the time). In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 111-120). ACM.

Hart, J., Ridley, C., Taher, F., Sas, C., & Dix, A. (2008, October). Exploring the facebook experience: a new approach to usability. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges* (pp. 471-474). ACM.

Hassenzahl, M. (2005). The thing and I: understanding the relationship between user and product. *Funology*, 31-42.

Hassenzahl, M. (2007). The hedonic/pragmatic model of user experience. *Towards a UX Manifesto*, 10-13.

Hassenzahl, M. (2008). User experience (UX): towards an experiential perspective on product quality. In *Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine* (pp. 11-15). ACM.

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. Behaviour & Information Technology, 25(2), 91-97.

Hertzum, M., & Jacobsen, N. E. (2001). The evaluator effect: A chilling fact about usability evaluation methods. *International Journal of Human-Computer Interaction*, *13*(4), 421-443.

Hole, L., & Williams, O. M. (2007). Gaining insight into the User eXperience. In *Proceedings of the'Towards a UX Manifesto workshop'at HCI*.

Holzinger, A. (2005). Usability engineering methods for software developers. *Communications of the ACM*, *48*(1), 71-74.

Iannella, R. (2009). Heuristic User Interface Evaluation: Three Case Studies of Dialog Design. In *Proceedings of OZCHI94-Australian CHISIG Annual Conference. Eds. Howard & Leung. Melbourne (Nov/Dec 1994)* (pp. 127-132).

International Organization for Standardization. (1998). *ISO 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs): Part 11: Guidance on Usability*.

Jeffries, R., & Desurvire, H. (1992). Usability testing vs. heuristic evaluation: was there a contest?. *ACM SIGCHI Bulletin*, *24*(4), 39-41.

Jeffries, R., Miller, J. R., Wharton, C., & Uyeda, K. (1991). User interface evaluation in the real world: a comparison of four techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 119-124). ACM.

Jokela, T. (2004). When good things happen to bad products: where are the benefits of usability in the consumer appliance market?. *Interactions*, *11*(6), 28-35.

Jordan, P. W. (1998). Human factors for pleasure in product use. *Applied ergonomics*, *29*(1), 25-33.

Kahn, M. J., & Prail, A. (1994). Formal usability inspections. *Usability inspection methods*, 141-171.

Kantner, L., & Rosenbaum, S. (1997). Usability studies of WWW sites: heuristic evaluation vs. laboratory testing. In *Proceedings of the 15th annual international conference on Computer documentation* (pp. 153-160). ACM.

Karapanos, E., Zimmerman, J., Forlizzi, J., & Martens, J. B. (2009). User experience over time: an initial framework. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 729-738). ACM.

Koca, A., Funk, M., Karapanos, E., Rozinat, A., van der Aalst, W. M., Corporaal, H., ... & Brombacher, A. C. (2009). Soft Reliability: an interdisciplinary approach with a user–system focus. *Quality and reliability engineering international*, *25*(1), 3-20.

Law, E. L. C., Roto, V., Hassenzahl, M., Vermeeren, A. P., & Kort, J. (2009). Understanding, scoping and defining user experience: a survey approach. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 719-728). ACM.

Law, E. L. C., Vermeeren, A. P., Hassenzahl, M., & Blythe, M. (2007, September). Towards a UX manifesto. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but not as we know it-Volume 2* (pp. 205-206). British Computer Society.

Ling, C., & Salvendy, G. (2005). Extension of heuristic evaluation method: a review and reappraisal. *Ergonomia IJE & HF*, *27*(3), 179-197.

Lyytinen, K., & Hirschheim, R. (1987) Information Systems Failures: A Survey and Classification of the Empirical Literature, Oxford Surveys *in Information Technology* (4) pp. 257-309.

Manning, H., McCarthy, J. C., & Souza, R. K. (1998). Why most web sites fail. *Interactive Technology Series*, *3*(7), 54-64.

Mantei, M. M., & Teorey, T. J. (1988). Cost/benefit analysis for incorporating human factors in the software lifecycle. *Communications of the ACM*, *31*(4), 428-439.

McNamara, N., & Kirakowski, J. (2006). Functionality, usability, and user experience: three areas of concern. *interactions*, *13*(6), 26-28.

Mendoza, V., & Novick, D. G. (2005). Usability over time. In *Proceedings of the 23rd annual international conference on Design of communication: documenting & designing for pervasive information* (pp. 151-158). ACM.

Nielsen, J. (1989). Usability engineering at a discount. In *Proceedings of the third international conference on human-computer interaction on Designing and using human-computer interfaces and knowledge based systems (2nd ed.)* (pp. 394-401). Elsevier Science Inc.

Nielsen, J. (1992). Finding usability problems through heuristic evaluation. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 373-380). ACM.

Nielsen, J. (1994a). Heuristic evaluation. *Usability inspection methods*, *17*, 25-62.

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

Nielsen, J. (1995). *Severity Ratings for Usability Problems,* Retrieved on 22.11.2013 from: http://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/

Nielsen, J. (1996). Putting the user in user-interface testing. *Software, IEEE,* *13*(3), 89-90.

Nielsen, J. (1999). Web research: Believe the data. *Retrieved* 2.12.2013 available at http://www.nngroup.com/articles/web-research-believe-the-data/

Nielsen, J. (2001). How to conduct a heuristic evaluation. *Retrieved* 6.2.2014 available at http://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/

Nielsen, J. (2003). Information foraging: Why Google makes people leave your site faster. *Jakob Nielsen's Alertbox*. *Retrieved* 29.11.2013 available at http://www.technical-expressions.com/layout/exercises/documents/Nielsen-InformationForaging-MarkedUp-Proxy-Align-Contrast.pdf

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 206-213). ACM.

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces, *Proc. ACM CHI'90 Conf.* pp: 249-256.

Oinas-Kukkonen, H., & Oinas-Kukkonen, H. (2013). *Humanizing the web: Change and social innovation*. Palgrave Macmillan.

Oliver, R. L. (2010). *Satisfaction: A behavioral perspective on the consumer*. ME Sharpe.

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction Design: Beyond Human-Computer Interaction.* John Wiley.

Rosenbaum, S., Rohn, J. A., & Humburg, J. (2000). A toolkit for strategic usability: results from workshops, panels, and surveys. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 337-344). ACM.

Savov, V. (2013). Hands-on with Jolla, the Nokia that could have been *What if Nokia hadn't abandoned MeeGo in favor of Windows Phone? Retrieved* 2.12.2013 *from* http://www.theverge.com/2013/11/29/5156446/jolla-smartphone-hands-on-preview

Spool, J., & Schroeder, W. (2001). Testing web sites: Five users is nowhere near enough. In *CHI'01 extended abstracts on Human factors in computing systems* (pp. 285-286). ACM.

Thong, J. Y., Hong, S. J., & Tam, K. Y. (2006). The effects of post-adoption beliefs on the expectation-confirmation model for information technology continuance. *International Journal of Human-Computer Studies*, *64*(9), 799-810.

Tuunanen, T., Myers, M., & Cassab, H. (2010). A Conceptual Framework for Consumer Information Systems Development. *Pacific Asia Journal Of The Association For Information Systems, 2*(1). Retrieved December 8, 2013, from http://journal.ecrc.nsysu.edu.tw/pajais/article/view/67/35

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: four longitudinal field studies. *Management science*, *46*(2), 186-204.

Vermeeren, A. P., Law, E. L. C., Roto, V., Obrist, M., Hoonhout, J., & Väänänen-Vainio-Mattila, K. (2010, October). User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries* (pp. 521-530). ACM.

Väänänen-Vainio-Mattila, K., & Wäljas, M. (2009). Development of evaluation heuristics for web service user experience. In *CHI'09 Extended Abstracts on Human Factors in Computing Systems* (pp. 3679-3684). ACM.

Weinschenk, S. (2005). Usability a business case. White Paper Human Factors International. Retrieved 28.11.2013 from: www.humanfactors.com

Wharton, C., Rieman, J., Lewis, C., & Polson, P. (1994). The cognitive walkthrough method: A practitioner's guide. In *Usability inspection methods* (pp. 105-140). John Wiley & Sons, Inc..

Wiklund, M. E. (2012). {Usability in Practice: How Companies Develop User-Friendly Products}.

Woolrych, A., & Cockton, G. (2001). Why and when five test users aren't enough. In *Proceedings of IHM-HCI 2001 conference* (Vol. 2, pp. 105-108). Toulouse,, France: Cépadèus.

Appendix A. Different Definitions for User Experience

All the aspects of how people use an interactive product: the way it feels in their hands, how well they understand how it works, how they feel about it while they're using it, how well it serves their purposes, and how well it fits into the entire context in which they are using it.

Alben, L. (1996). Defining the criteria for effective interaction design.*interactions*, *3*(3), 11-15.

All aspects of the end-user's interaction with the company, its services, and its products. The first requirement for an exemplary user experience is to meet the exact needs of the customer, without fuss or bother. Next comes simplicity and elegance that produce products that are a joy to own, a joy to use. True user experience goes far beyond giving customers what they say they want, or providing checklist features. In order to achieve high-quality user experience in a company's offerings there must be a seamless merging of the services of multiple disciplines, including engineering, marketing, graphical and industrial design, and interface design.

http://www.nngroup.com/articles/definition-user-experience/

The overall experience, in general or specifics, a user, customer, or audience member has with a product, service, or event. In the Usability field, this experience is usually defined in terms of ease-of-use. However, the experience encompasses more than merely function and flow, but the understanding compiled through all of the senses.

http://www.nathan.com/ed/glossary/index.html

Every aspect of the user's interaction with a product, service, or company that make up the user's perceptions of the whole. User experience design as a discipline is concerned with all the elements that together make up that interface, including layout, visual design, text, brand, sound, and interaction. UE works to coordinate these elements to allow for the best possible interaction by users.

http://www.usabilitybok.org/glossary

User experience (UX) is about how a person feels about using a system. User experience highlights the experiential, affective, meaningful and valuable aspects of human-computer interaction (HCI) and product ownership, but it also covers a person's perceptions of the practical aspects such as utility, ease of use and efficiency of the system. User experience is subjective in nature, because it is about an individual's performance, feelings and thoughts about the system. User experience is dynamic, because it changes over time as the circumstances change.

http://en.wikipedia.org/wiki/User_experience

User Experience (abbreviated: UX) is the quality of experience a person has when interacting with a specific design.

Knemeyer, Dirk and Svoboda, Eric (2006). User Experience - UX. Retrieved 16 March 2014 from http://www.interaction-design.org/encyclopedia/user_experience_or_ux.html

A result of motivated action in a certain context. User's previous experiences and expectations influence the present experience; this present experience leads to more experiences and modified expectations. Mäkelä & Fulton Suri (2001)

Mäkelä, A., & Fulton Suri, J. (2001, June). Supporting users' creativity: Design to induce pleasurable experiences. In *Proceedings of the International Conference on Affective Human Factors Design* (pp. 387-394). Asean Academic Press, London.

A consequence of a user's internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (e.g. complexity, purpose, usability, functionality, etc.) and the context (or the environment) within which the interaction occurs (e.g. organisational/social setting, meaningfulness of the activity, voluntariness of use, etc.).

Hassenzahl, M., & Tractinsky, N. (2006). User experience - a research agenda. Behaviour & Information Technology, 25(2), 91-97.

The value derived from interaction(s) [or anticipated interaction(s)] with a product or service and the supporting cast in the context of use (e.g., time, location, and user disposition).

Sward, D., & Macarthur, G. (2007, September). Making user experience a business strategy. In *E. Law et al.(eds.), Proceedings of the Workshop on Towards a UX Manifesto* (Vol. 3, pp. 35-40).

The user experience considers the wider relationship between the product and the user in order to investigate the individual's personal experience of using it.

McNamara, N., & Kirakowski, J. (2006). Functionality, usability, and user experience: three areas of concern. *interactions*, *13*(6), 26-28.

Users' perceptions of interaction that constitute qualities of use.

Colbert, M. (2005). User experience of communication before and during rendezvous: interim results. *Personal and Ubiquitous Computing*, *9*(3), 134-141.

An activity of encounter by a computer user with the auditory and visual presentation of a collection of computer programs. It is important to note that this includes only what the user perceives and not all that is presented.

http://msdn.microsoft.com/en-us/library/bb246417.aspx

The entire set of affects that is elicited by the interaction between a user and a product, including the degree to which all our senses are gratified (aesthetic experience), the meanigs we attach to the product (experience of meaning), and the feelings and emotions that are elicited (emotional experience).

Hekkert, P. (2006). Design aesthetics: principles of pleasure in design.*Psychology science*, *48*(2), 157.

UX is a momentary, primarily evaluative feeling (good-bad) while interacting with a product or service.

Hassenzahl, M. (2008, September). User experience (UX): towards an experiential perspective on product quality. In *Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine* (pp. 11-15). ACM.

A person's perceptions and responses that result from the use or anticipated use of a product, system or service. ISO 9241-210 (2010)

A set of material rendered by a user agent which may be perceived by a user and with which interaction may be possible.

http://www.w3.org/TR/di-gloss/#def-user-experience

Encompasses all aspects of a digital product that users experience directly—and perceive, learn, and use—including its form, behavior, and content. Learnability, usability, usefulness, and aesthetic appeal are key factors in users' experience of a product.

http://www.uxmatters.com/glossary/

The user experience, mostly called "customer experience" when referring to e-commerce websites; the totality of the experience of a user when visiting a website. Their impressions and feelings. Whether they're successful. Whether they enjoy themselves. Whether they feel like coming back again. The extent to which they encounter problems, confusions, and bugs.

http://www.usabilityfirst.com/glossary/customer-experience/

User experience = Convenience + Design – Cost. Convenience is the king. What makes a product convenient is quite often what makes it usable. It might also relate to the availability of the product. It might also have something to do with laziness and productivity. Defining "convenience" is by no means an easy task. As is with everything else in this chart, convenience is subjective. Design is what makes a product liked and attractive, even before it has been used. Design is what makes you want the product. It is beauty, the touch of a famous designer, a likable company, character—pretty much what brand value is thought to be.

http://www.nnyman.com/personal/2005/11/18/the-user-experience-equation/

The user experience is the totality of end-users' perceptions as they interact with a product or service. These perceptions include effectiveness (how good is the result?), efficiency (how fast or cheap is it?), emotional satisfaction (how good does it feel?), and the quality of the relationship with the entity that created the product or service (what expectations does it create for subsequent interactions?). Kuniavsky (2010)

Kuniavsky, M. (2010). *Smart Things: Ubiquitous Computing User Experience Design: Ubiquitous Computing User Experience Design*. Elsevier.

The overall experience and satisfaction a user has when using a product or system. Old Wikipedia definition, still used e.g. at

http://www.bitpipe.com/tlist/User-Experience.html

The overall perception and comprehensive interaction an individual has with a company, service or product. A positive user experience is an end-user's successful and streamlined completion of a desired task.

Goto, K. (2004). Brand value and the user experience. *Digital Web Magazine*.

UX   =   the   sum   of   a   series   of   interactions
User experience (UX) represents the perception left in someone's mind following a series of interactions between people, devices, and events – or any combination thereof.

http://www.fatdux.com/blog/2009/01/10/a-definition-of-user-experience/

User experience stands for the quality of a global experience as perceived by a person (user) interacting with a system.

http://www.use-design.com/en/user-experience

Users' judgement of product quality arising from their experience of interaction, and the product qualities which engender effective use and pleasure.

Sutcliffe, A. (2009). Designing for user engagement: Aesthetic and attractive user interfaces. *Synthesis lectures on human-centered informatics*, *2*(1), 1-55.

Appendix B. UX Heuristic evaluation form

UX Heuristic form by Arhippainen (2013)

UX Heuristic	Y/ N /NA*	Comment
1. Ensure usability		
2. Provide utility matching with the user's values		
3. Surpass the user's expectations		
4. Respect the user		
5. Design the product or service to fit the intended contexts		
6. Provide several ways to interact, leave choice for the user		
7. Respect the user's privacy and security		
8. Support the user's activities - do not force		
9. Go for a perfect visual design		
10. Give a surprise gift		

*Y= Yes   N= No   NA= Not Applicable