

# **The Influence of the Usability of a Reputation System on its Utility**

**Maysa Jamil**

Kongens Lyngby 2012

IMM-M.Sc.-2012-136

Technical University of Denmark  
Informatics and Mathematical Modelling  
Building 321, DK-2800 Kongens Lyngby, Denmark  
Phone +45 45253351, Fax +45 45882673  
reception@imm.dtu.dk  
www.imm.dtu.dk IMM- M.Sc. -2012-136



# Abstract

---

Usability, or the “ease of use” of a website, is important in order to allow for a user to carry out the intended task in an efficient and effective manner. For website that uses reputation system, as these chosen in this thesis, such aspect is equally important. But in order to ensure a system which is “Usable” there must be carried out usability activities to that system.

The goal of the thesis was to carry out a usability study on known websites that use reputation system in order to analyze the current level of usability on these websites. By indentifying the level of the usability of these websites it should be possible to present recommendations in order to improve usability in future reputation-based websites.

In the thesis there was planned a framework for carrying out the usability study. The proposed framework was based upon various methods available on the field of Human Computer Interaction (HCI).

The methods of choice to analyze the level of the usability of the chosen websites were heuristic evaluation based on Nielsen’s heuristics; where each of the heuristics was applied to each website’s reputation system. The second method was the think aloud test; where six test participants have been chosen to do some test tasks, and thereby making it possible to indentify its usability strengths and weaknesses from expert and users point of view.

Based on the usability evaluation we chose three websites that have (the best, natural and worst) usability to conduct another usability study to examine whether the usability has an impact on the overall utility of the reputation system.

The usability methods of choice to analyze the impact the usability of the reputation system on its utility were the interviews and survey.

After the conducted usability approach we found that the level of usability has an influence on how the accessible, utilizable and understandable the reputation information is to the user

On basis of our findings, we presented some recommendations that should be considered when integrating the reputation systems in the websites.

# Acknowledgment

---

This thesis was submitted as a part of the requirements for completing the Master degree program at the Technical University of Denmark (DTU). The Master thesis was carried out in the period April until September 2012.

I would like to express my sincere gratitude to my supervisor at DTU, Associate Professor Christian D. Jensen, for his guidance, feedback and support throughout this project.

# Table of contents

---

1. Introduction .....	1
1.1 Overview .....	1
1.2 Aim of Study .....	2
1.3 Research Questions .....	2
1.4 Study Activities .....	2
1.5 Outline of the Thesis.....	2
Chapter 2.....	4
2. Background for Reputation Systems and Usability.....	4
2.1 Reputation Systems .....	4
2.2 General model and examples.....	4
2.2.1 Centralized Reputation Systems.....	5
2.2.2 Distributed Reputation Systems .....	6
2.3 Reputation Computation Engines.....	7
2.4 The Usability.....	8
2.5 RS Usability.....	10
Chapter 3.....	11
3. The Research Framework & Methodology.....	11
3.1 The framework (approach) of the study.....	11
3.2 The chosen sites .....	12
3.3 The Users Selection .....	13
3.4 The chosen usability methods.....	13
3.4.1 Heuristic Evaluations for Nielsen.....	14
3.4.2 Nielsen Heuristics for Reputation Systems.....	16
3.4.3 Think aloud testing.....	26
3.4.4 Survey and Interview .....	29
3.5 Reliability and Validity .....	29
3.5.1 Reliability.....	30
3.5.2 Validity.....	30
Chapter 4.....	31

4.	Data Presentation .....	31
4.1	Application of Nielsen Ten Heuristics .....	31
4.1.1	eBay.com .....	31
4.1.2	Amazon.com .....	34
4.1.3	Yelp.com .....	39
4.1.4	Trustpilot.com.....	43
4.1.5	Epinions.com .....	48
4.1.6	iTunes Apps Store .....	53
4.1.7	Youtube.com .....	59
4.2	Applications of Think aloud and Remote Testing.....	66
4.2.1	eBay.com .....	66
4.2.2	Amazon.com .....	66
4.2.3	Yelp.com .....	67
4.2.4	TrustPilot.com.....	68
4.2.5	Epinions.com .....	69
4.2.6	iTunes Apps Store .....	70
4.2.7	YouTube.com .....	71
4.3	Usability of the reputation system of the chosen Websites.....	72
4.3.1	eBay.....	72
4.3.2	Amazon.....	73
4.3.3	Yelp.....	74
4.3.4	Trustpilot .....	74
4.3.5	Epinions .....	75
4.3.6	iTunes Apps Store .....	76
4.3.7	YouTube.....	77
4.4	Applications of Survey and Interviews.....	79
Chapter 5	.....	83
5.	Discussion and Recommendations.....	83
5.1	RQ1. Does the most common websites that use some kind of reputation system have a high level of usability?.....	83
5.1.1	Within heuristic evaluation .....	83
5.1.2	Within think aloud testing.....	84
5.2	RQ2. Does the reputation information representation have an impact on users?.....	85
5.2.1	Within survey and interviews .....	85

5.3 RQ3. Does the usability of the reputation system have an impact on the user review and decision making?.....	86
5.3.1 Within survey and interviews .....	86
Chapter 6 .....	88
6. Conclusion.....	88
Chapter 7 .....	89
7. Bibliography .....	89
Appendix .....	91
Appendix A. Usability Test Script .....	91
1. Test facilitator: .....	91
2. Pre-test interview questions.....	91
3. Test tasks: .....	91
4. Interview after test.....	93



# Chapter 1

## 1. Introduction

---

This chapter provides a background to the study topic, the goal of the thesis, research questions and the study activities. The final section of this chapter provides a brief overview of the thesis structure.

### 1.1 Overview

There are several reputation-based systems available online. These reputation systems are a key success factor of many websites. The main goals for using reputation systems are:

- To help users and consumers to have better understanding of the information, products and services being provided, i.e. to provide information to help assess whether an entity is trustworthy (trust assessment).
- To encourage entities to behave in a trustworthy manner, i.e. to encourage good behavior. [17]
- To discourage less reliable entities from participating in an interaction, i.e. to make chiseling and cheating rare and losing propositions. [18]

The usability is a quality attribute that assesses how easy user interfaces are to use. So the usability of the reputation system is very important to help the user understand the system and to know how to use it, i.e. the utility (the system functionality).

Usability and utility are equally important and together determine whether something is useful: It's no good if the system is easy, but doesn't do what you want. It's also bad if the system can do what you want, but you can't use it, because it's too difficult to use it.

- Definition: **Utility** = whether it provides the features you need.
- Definition: **Usability** = how easy & pleasant these features are to use.
- Definition: **Useful** = usability + utility.

The utility of the reputation system are decision making and review the service or the provider, and if the users cannot find the reputation information or cannot understand it, the system are useless. [19]

## **1.2 Aim of Study**

The main goal of the thesis is to carry out a usability study on reputation systems in order to identify the level of the usability of those reputation systems. Furthermore, it does this with the intention of being able to identify key factors that will influence on the utility of the reputation systems. It should be possible to present recommendations in order to improve the usability of the reputation systems in the future.

## **1.3 Research Questions**

Based on our aim of the study in the section 1.2, we have formulated three research questions they are as follows:

**RQ1.** Does the most common websites that use some kind of reputation system have a high level of usability?

**RQ2.** Does the reputation information representation have an impact on users?

**RQ3.** Does the usability of the reputation system have an impact on the user review and decision making?

## **1.4 Study Activities**

By addressing the research questions, a framework (approach) will be planned for carrying out the usability study of the reputation-based systems. The framework will be based on recognized usability evaluation methods in the field of Human Computer Interaction (HCI). The websites chosen for this study are well known websites and have popularity each in their field. As part of the framework and the study recommendation will be presented for future reputation-based systems development.

## **1.5 Outline of the Thesis**

This thesis is divided in six chapters. Chapter one gives an introduction to the study topic, research questions and the study activities. Chapter two will present the background for the reputation system and the usability. The third chapter presents the research framework and methodology and the choice of the usability methods to carry out the

study. Chapter four will present the collected data from the research methods. In chapter five we will discuss the collected data in contrast with the research questions. We will also make some recommendations. Finally in chapter six we will make the conclusion of the study.

# Chapter 2

## 2. Background for Reputation Systems and Usability

---

This chapter will present the general model of the reputation system and an introduction to the usability and its attributes.

### 2.1 Reputation Systems

We will define reputation according to the Concise Oxford dictionary:

*“Reputation is what is generally said or believed about a person’s or thing’s character or standing”.*

Reputation can be considered as a collective measure of trustworthiness based on the referrals or ratings from members in a community. In decision making, personal experience, i.e., trust, usually carries more weight than referrals, i.e., reputation. However, if direct experiences are lacking, decision making has to be based on referrals from others.

Reputation systems are widely used in different aspects in our life; in electronic commerce, social network, search engine, and so on. A reputation system collects, distributes, and aggregates feedback about participants past behavior. Though few of the producers or consumers of the ratings know each other, these systems help people decide whom to trust, encourage trustworthy behavior, and show people who are unskilled or dishonest. [1, 13, 14]

### 2.2 General model and examples

In this section we will describe the reputation systems types and how the ratings and reputation scores are communicated between participants in a reputation system. There are two types of reputation systems, they are centralized and distributed systems.

## 2.2.1 Centralized Reputation Systems

In the centralized reputation systems, a reputation center will collect all the information about the performance of a community participant. This information, e.g. in the form of ratings, is collected from other community members who have had direct experience with that participant.

The central reputation center that collects all the ratings typically derives a reputation score for each participant, and makes all scores publicly available. The participants can then use each other's scores, for example in decision making. The idea is that transactions with reputable participants are likely to have more favorable outcomes than transactions with disreputable participants. Figure 1 shows a general centralized reputation system, where C and P represent the transaction partners with a history of transactions in the past, and they consider transacting with each other in the present. Fig 1.b shows a present transaction depending on the experience of the other transitions in the past fig 1.a.

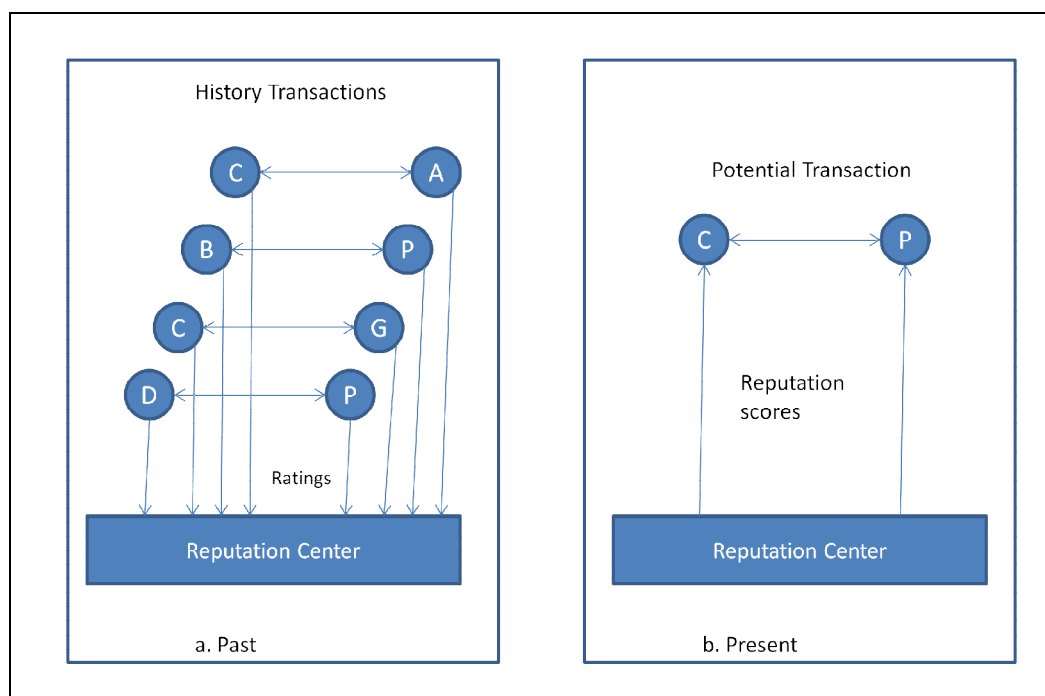


Figure 1: General framework for a centralized system [1]

After a transaction is completed, the partners provide ratings about each other's performance in that transaction. The reputation centre collects ratings from all the partners, and continuously updates each partner's reputation score as a function of the received ratings. The Updated reputation scores are provided online for all the partners to see, and can be used by other partners to help them to decide whether or not to deal with that partner.

The two fundamental aspects of centralized reputation systems are:

1. Centralized communication protocols that allow participants to provide ratings about transaction partners to the central authority, as well as to obtain reputation scores of potential transaction partners from the central authority.
2. A reputation computation engine used by the central authority to derive reputation scores for each participant, based on received ratings, and possibly also on other information. Will discuss it in the section Reputation Computation Engines. [1]

## 2.2.2 Distributed Reputation Systems

In the distributed reputation system there is no central reputation center for collecting ratings or obtaining reputation scores of others. Instead, there can be distributed stores where ratings can be collected, or each member records the opinion about each experience with the target party, and provides this information on request from other relying members, who consider transacting with that target party. These people should themselves find the distributed stores, or try to obtain ratings from as many members as possible who have had direct experience with that target party. Figure 2 shows this.

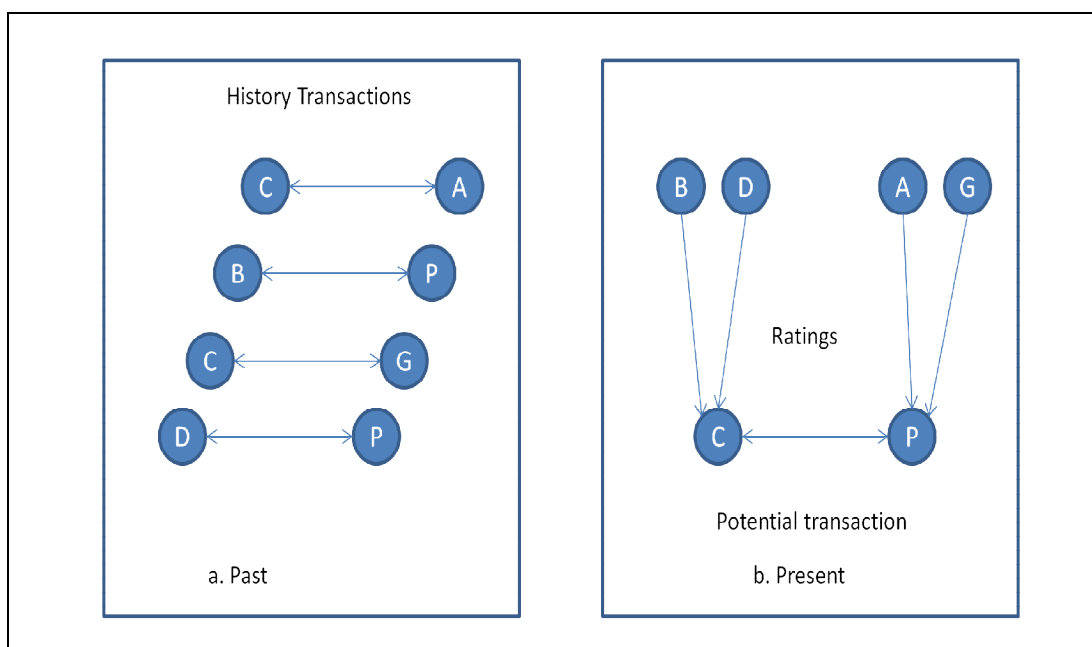


Figure 2: General framework for a distributed reputation system [1]

Then the relying members compute the reputation score based on the received ratings. In case the relying member has already had direct experience with the target party, the

experience from that transaction can be taken into account as private information, possibly carrying a higher weight than the received ratings.

The two fundamental aspects of distributed reputation systems are:

1. A distributed communication protocol that allows participants to obtain ratings from other members in the community.
2. A reputation computation method used by each individual agent to derive reputation scores of target parties based on received ratings, and possibly on other information. This will presents in section Reputation Computation Engines. [1]

## 2.3 Reputation Computation Engines

The reputation scores can be calculated based on own experience, on others experience, or on a combination of both. Reputation systems are typically based on public information in order to reflect the community's opinion in general. But some systems take both public and private information as input. The private information is the resulting from personal experience, which is normally considered more reliable than public information, such as ratings from other parties. There are many ways to compute the reputation measures, the most popular methods are:

### 1. Simple Summation or average of Ratings

This method is the simplest form of computing reputation scores. It sums the number of positive ratings and negative ratings separately, and to keep a total score as the positive score minus the negative score. This is the principle used in eBay's and Amazon's reputation systems. The advantage is that anyone can understand this principle.

### 2. Bayesian Systems

In this method the binary ratings are taken as input (i.e. positive or negative), and are based on computing reputation scores by statistical updating of beta probability density functions (PDF). The updated reputation score is computed by combining the previous reputation score with the new rating. The advantage of Bayesian systems is that they provide a theoretically sound basis for computing reputation scores. This principle is used in Yelp. [1]

## 2.4 The Usability

If a website is difficult to use, people leave. If the homepage fails to clearly state what a company offers and what users can do on the site, people leave. If users get lost on a website, they leave. If a website's information is hard to read or doesn't answer users' key questions, they leave. There's no such thing as a user reading a website manual or otherwise spending much time trying to figure out an interface. There are plenty of other websites available; leaving is the first line of defense when users encounter a difficulty. [19]

The system to be socially acceptable, it should be practical acceptable within various categories such as cost, support, reliability, etc., as well as the category of usefulness. The usefulness is the issue of whether the system can be used to achieve some desired goal. The usefulness can be divided into two categories of Utility and Usability, where the utility is the question of whether the functionality of the system in principle can do what is needed, and usability is the question of how well users can use that functionality. Figure 3 shows the simple model of system acceptability. So the system to be useful it should be both usable and has utility.

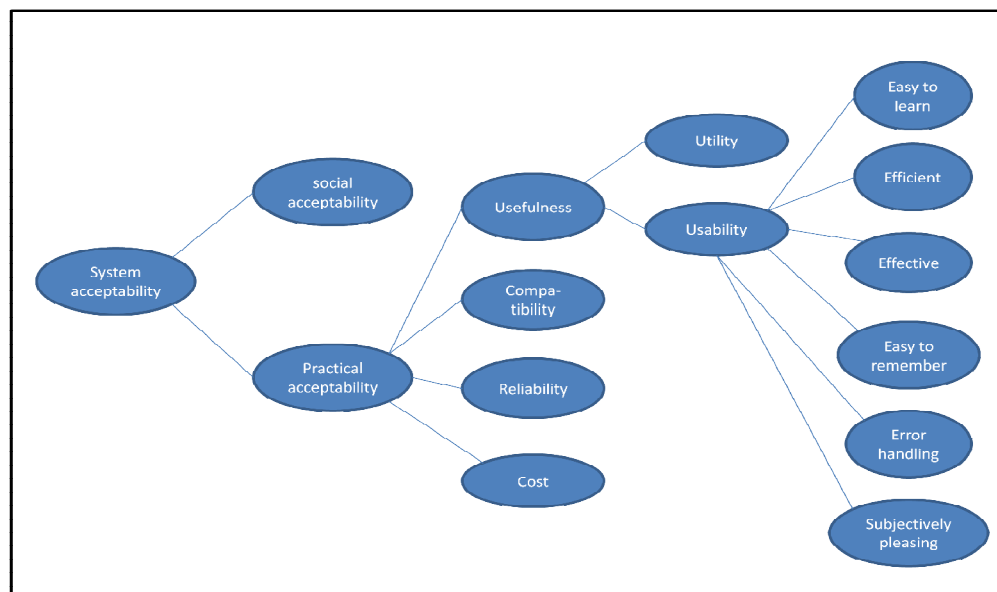


Figure 3 the simple model of system acceptability [2]

The usability has many definitions but the most two important definitions are:  
The first one presented by the ISO standard 9241-11:

*"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."*



And the second usability definition is presented by one of the pioneers within the field of usability: Jakob Nielsen:

*"Usability is a quality attribute that assesses how easy user interfaces are to use."*

The usability is not a single, one-dimensional property of a user interface. It has multiple components according to Nielsen [2] and is traditionally associated with these five usability attributes:

**Learnability:** the system should be easy to learn so that the user can rapidly start getting some work done with the system. It is the most fundamental usability attribute, since most systems need to be easy to learn, and since the first experience most people have with a new system is that of learning to use it. The users should be able to complete a certain task successfully in a certain minimum time.

**Efficiency:** the system should be efficient to use, so that once the user learns the system, a high level of productivity is possible. The experience is the key to measure efficiency, so to measure the efficiency we need experienced users. The users are considered experienced if they have been users for more than a certain amount of time, such as months or a year. A typical way to measure efficiency of use is to bring experienced test users and measure the time it takes these users to perform some specific tasks.

**Memorability:** the system should be easy to remember, so that the casual user is able to return to the system after some period of not having used it, without having to learn everything all over again.

**Errors:** the system should have a low error rate, so that users make few errors during the use of the system, and so that if they do make errors they can easily recover from them. Further, catastrophic errors must not occur. Typically, an error is defined as any action that does not accomplish the desired goal. Some error can be corrected right away by the system, like these errors will just slow down the transaction time but don't count as catastrophic errors, which lead to a faulty work, or destroy the user's work, making them difficult to recover from.

**Satisfaction:** the system should be pleasant to use, so that users are subjectively satisfied when using it; they like it. It is the comfort and acceptability of the work system to its users and other people affected by its use. [2, 15]

In this thesis we will use Nielsen's definition of usability, with its impact on the system functionality (utility). One can evaluate the system's functionality by adopting the same methods used for evaluating usability according to Nielsen. The goal of the usability effort

in this thesis was to find the strengths and weaknesses of the reputation system in the common websites, and to find its impact on the functionality (utility) of the reputation system.

## **2.5 RS Usability**

So much money is being spent on reputation-based websites. Yet so many websites don't seem to have considered the usability of their reputation system, resulting in users giving up using these sites. There are many guidelines that should be considered in each reputation system, we discussed these guidelines in section 3.4.2.

# Chapter 3

## 3. The Research Framework & Methodology

---

In this chapter we will describe the chosen sites and the different usability methods used in our study and explain the goals behind each method. We will illustrate the tasks that have been conducted in our study.

### 3.1 The framework (approach) of the study

The goal of the thesis was to carry out a usability study on Reputation systems in order to identify the level of the usability of the reputation systems and what had a positive and negative impact on the utility of the Reputation systems. It should be possible to present recommendations in order to improve the usability of the reputation systems in the future.

A framework (approach) as shown in figure 4 will be proposed for carrying out the usability study. The framework will be based on recognized usability evaluation methods.

The approach consists of these steps:

1. Define the project goals, select the sites with a reputation system to be measured, decide the extent of the usability work needed, and identify the resources needed to carry out the usability study.
2. Know the user; it's important to identify the user who will use the system and to determine the skill level.
3. Select the usability evaluation methods that will be used, and evaluate the usability of the system by using these chosen methods.
4. After finding the result of the evaluation, present recommendations for future development.

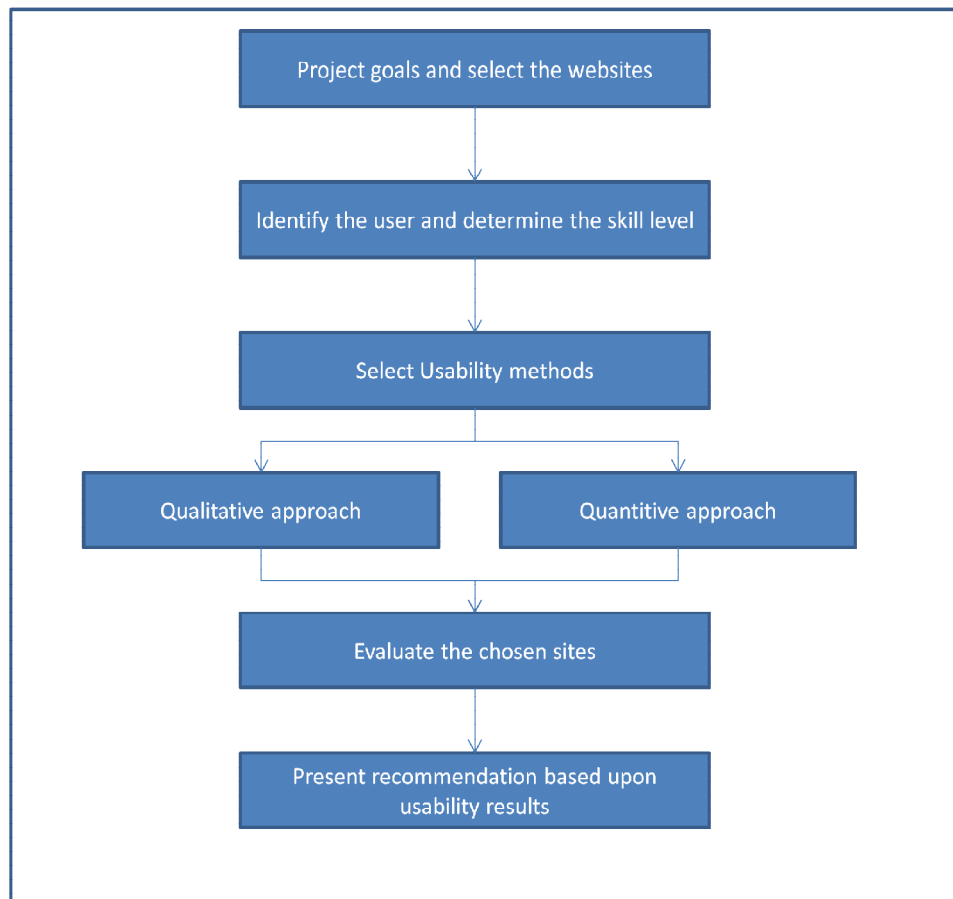


Figure 4 Framework for the study

### 3.2 The chosen sites

We have chosen seven websites to evaluate their usability. Each website is popular in its category (shopping, video sharing etc.) and they are mainly employing reputation system:

- eBay.com (shopping)
- Amazon.com (shopping)
- Yelp.com (reviews and recommendation)
- Trustpilot.com (reviews)
- Epinions.com (reviews and recommendation)
- apple.com/iphone/from-the-app-store (Smartphone apps store)
- Youtube.com (video sharing)

### 3.3 The Users Selection

Before we conduct our research, we think about our sample selection. So we have selected users depending on our study needs. The selection criteria are:

- We have selected users with at least bachelor level of study and variety of educational background, from different countries; therefore we can get benefit from their knowledge and different way of thinking.
- We have selected both male and female users with different ages between 22-60 years; therefore we can get benefit from the young people who are more adventurous with using the reputation-based sites and the senior people who are more cautious with the reputation-based sites.
- For the think aloud test we have selected six experienced users (in internet), however four of them have used the chosen websites; therefore we can get more benefit from their experience in these websites. We can also get benefit from the other two users who didn't use the chosen websites, because they are not biased to a specific website.
- For the think aloud test we have selected six users who have been tested before in a usability test; therefore we can get benefit of their usability experience.
- For the interview and the survey we have selected 48 users who have experience with the three filtered websites, so we can get benefit from the users' neutrality and their experience in these websites.

### 3.4 The chosen usability methods

There are two choices of research approaches available: qualitative and quantitative approach, so to choose which one of them to use, we should first consider our research questions and see which approach is suitable to our work.

#### **Qualitative research:**

Is the type of the research that provides answers to our questions, produces deeper understanding to a given research problem and collects evidence. This research provides descriptions of how people experience a given research issue. The qualitative methods are typically flexible, for example the qualitative methods ask mostly "open-ended"

questions that are not necessarily worded in exactly the same way with each participant. With open-ended questions, participants are free to respond in their own words. The aim of a qualitative research is to understand, not to explain. [3]

### **Quantitative Research:**

Is the type of the research that measure and analyze relationships between entities. The quantitative methods are typically inflexible, for example the quantitative methods ask mostly "closed-ended" or fixed questions that are identical to all the participants. The aim with this research is to explain and predict about future. [3]

As the aim of our study is to evaluate the usability level on the reputation system from the expert and users point of view and to find any flaws in the reputations systems and to explain the influence of the usability of the reputation system on its utility, therefore we find that it's best to use both qualitative and quantitative approaches in our research.

There are many methods available in the field of Human Computer Interaction (HCI) that can help us to evaluate the level of the usability of the reputations systems to websites. Heuristic Evaluation, Think Aloud and Remote Testing are the chosen method to evaluate and test the usability of the reputation information of the chosen sites. The interviews and survey are the chosen method to gather information what users need, behave, take decision, like or dislike about these websites and to find if the usability of the reputation system has an influence of its utility.

#### **3.4.1 Heuristic Evaluations for Nielsen**

Heuristic evaluation is the most popular method of the usability inspection methods where usability specialists or evaluator examine the website and judge its compliance with recognized usability principles (heuristics). A "double" specialist, that is, someone who is an expert in usability principles or human factors as well as an expert in the domain area (such as reputation system, financial services, and so on, depending on the application), or in the particular technology employed by the product, can be more effective than one without such knowledge. [16, 21]

The main goal is to evaluate the level of the usability in the most common reputation-based sites and how well the ten heuristic evaluation of Nielsen works for these reputation-based sites. We will focus on the reputation information and whether if it works or fails and how we can improve it.

These are ten general principles for user interface design:

1. **Visibility of system status**

The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

2. **Match between system and the real world**

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

3. **User control and freedom**

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

4. **Consistency and standards**

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

5. **Error prevention**

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

6. **Recognition rather than recall**

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

7. **Flexibility and efficiency of use**

Accelerators -- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

8. **Aesthetic and minimalist design**

Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

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

Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

10. **Help and documentation**

Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large. [16]

### 3.4.2 **Nielsen Heuristics for Reputation Systems**

In this section we will analyze the meaning of the Nielsen heuristics in the context of reputation systems.

#### 1. **Visibility of system status**

The visibility of the system status should be presented in a clear manner allowing for the user immediately identify the system status.

The most important things the users need to know when visit reputation systems are:

- Where am I? The system should show a clear indication to where the user is, so he can see what categories and what sorting options he selects and in which page he is, for example figure 5 shows clearly the account name, and the name and address of the restaurant the user wants to review.
- What is going on? The system should keeps users informed what's going on and when the user interacts with the system the changes should be made clearly, and the updates should be made within moments of submission. The user needs some sort of feedback that lets him immediately know that his action is being processed. Figure 6 shows the system responds within a few seconds after the user left a feedback to a seller.



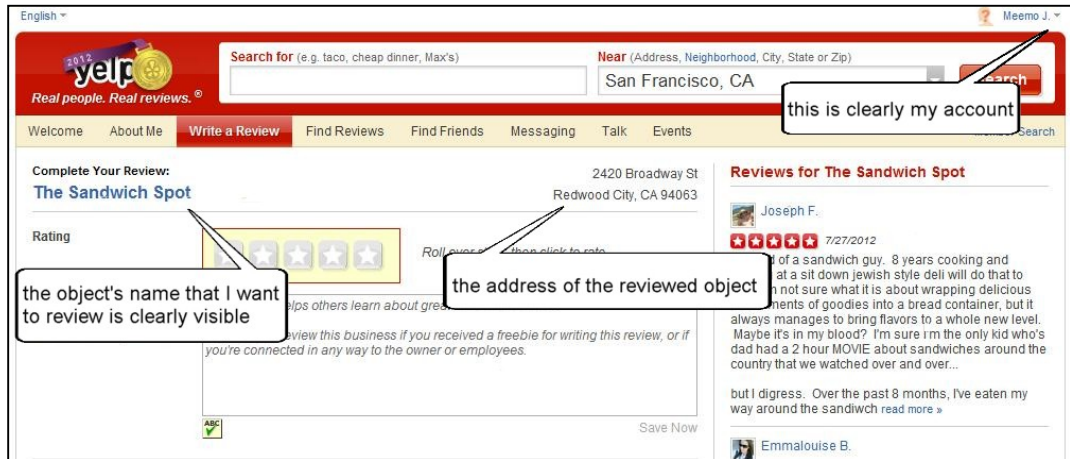


Figure 5: view showing it's me and the name & address of the place that I want to review (yelp.com)

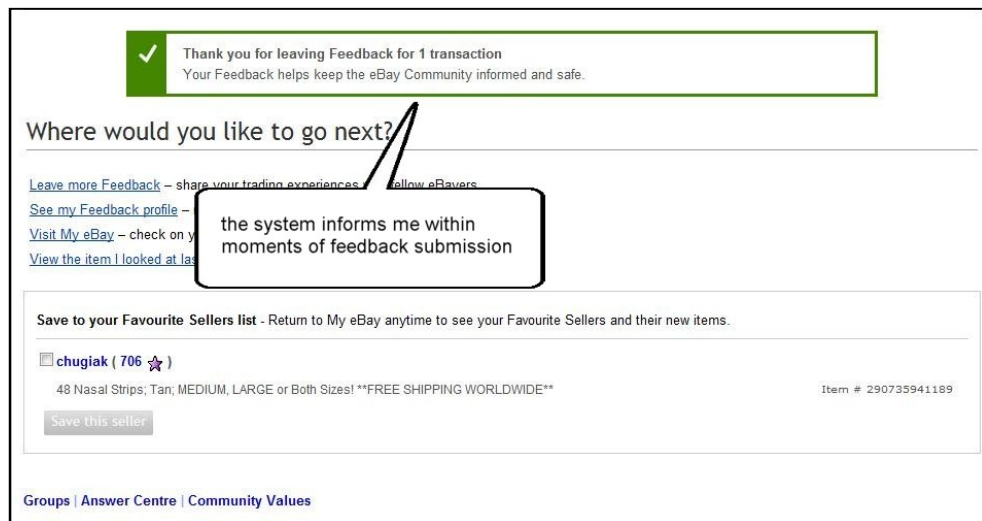


Figure 6: view showing the system responds within moments of feedback submission (eBay.com)

## 2. Match between system and the real world

To make it easy for the users to work with the reputation systems, the reputation categories and information should be in plain language (the user language) and followed real-world conventions as shown in figure 7, and the information should appear in a logical order, one can notice in figure 8 how the "Member quick links" are not alphabetical order but they are well grouped.



Figure 7: view showing the reputation information in plain language (eBay.com)

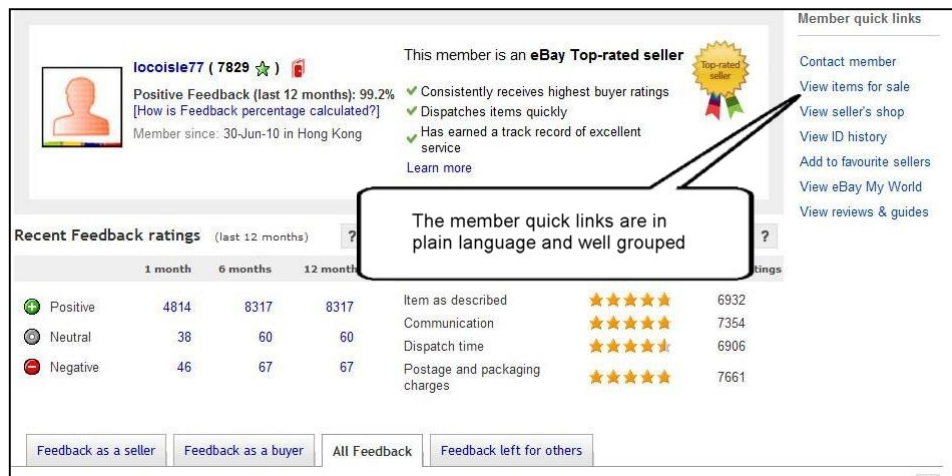


Figure 8: view showing the well grouped categories (eBay.com)

### 3. User control and freedom

User control and freedom can have a significant impact on user satisfaction. When the users have an impression that they aren't in control of the system then it might be a cause for user frustration and subsequently have a negative impact on the users' attitude towards the use of the system. So the system should provide a number of areas that can be edited, modified or deleted. The user should ensure that the system meets the control requirements of both novice and experienced users. An example of user control and freedom might be when I change my mind about an old review that I wrote, there should be always a way to go back to edit or remove that review, as shown in figure 9.



Figure 9: view showing the possibility to change my old review (yelp.com)

#### 4. Consistency and standards

The user should understand how the reputation ratings are calculated and what they mean. If the users don't understand what the star rating refers to or what function any icon or link carries out, the users will hesitate to use it. It should be easy for the user to find the information or specification he needs, and should know where to go if he needs to and what types of information to expect to find when he do an action, e.g. when a user want to contact a member, he should go to a label "contact member", rather than some obscure reference. Figures 10 and 11 show the consistency and standards should be available in each reputation system.

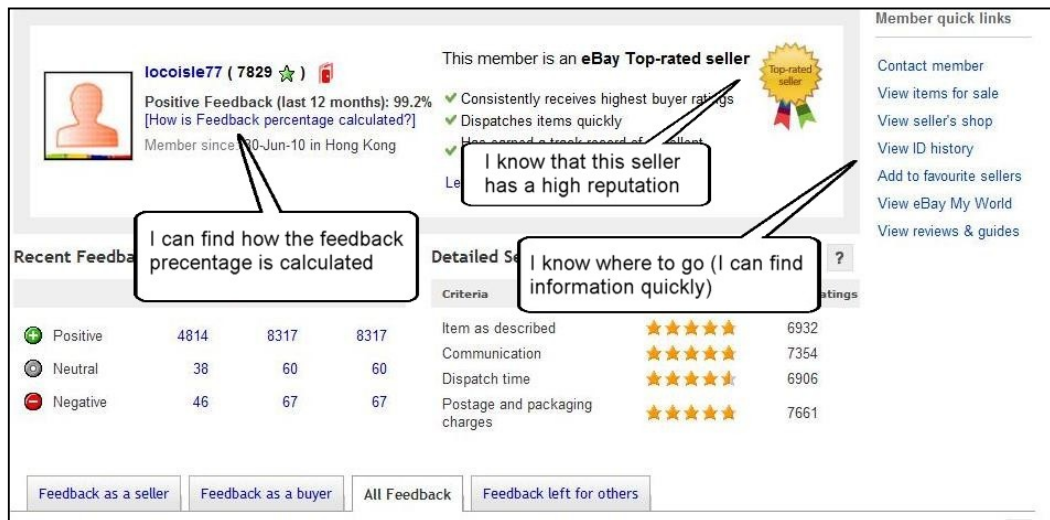


Figure 10: view showing a typical seller reputation (eBay.com)

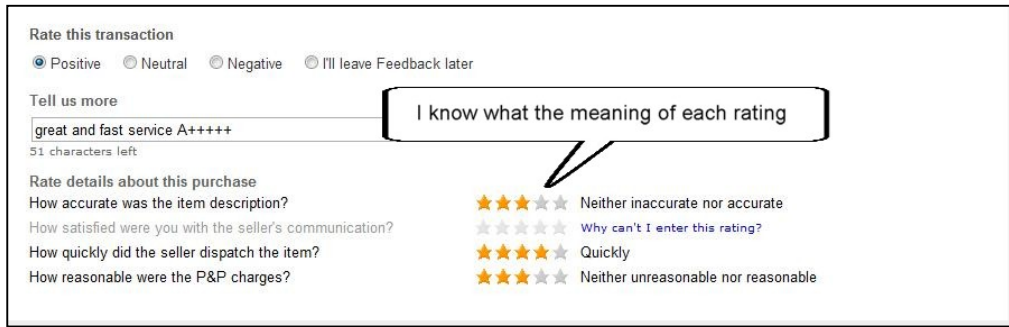


Figure 11: view showing the ease of understanding the reputation ratings(eBay.com)

## 5. Error prevention

There are many places where error can occur on the reputation systems, and the user may conduct these errors. The ideal system should prevent a problem from occurring in the first place, but if the errors do occur, the system should provide a user friendly message in plain language rather than codes, e.g. if the user wants to leave an opinion sentence about a company, he forgets to write at least a sentence with three words, a user friendly message will appear, figure 12.

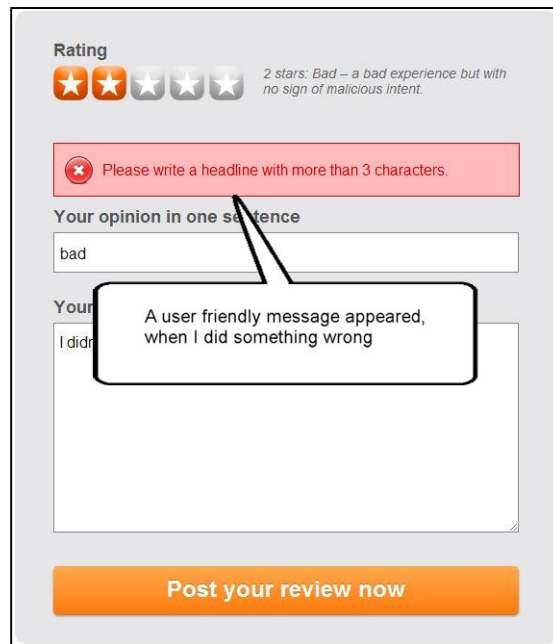


Figure 12: view showing a user friendly message when an error do occur (trustpilot.com)

## 6. Recognition rather than recall

The system's objects, actions, options, and links should be visible so the user's memory load would be minimized. The system's information should be visible or easily retrievable whenever needed. The system should have a history of all the user's previous

transactions and reviews. The user should retrieve any review or old orders, so this will help him to remember exactly what he had ordered years ago, e.g. the user can retrieve what users recently review a seller, as shown in figure 13. The user should have the possibility to take a look at the transition's details before he rates the seller, because maybe after at least two weeks the user doesn't remember what the details of this order are, figure 14. The user should also have the possibility to save the reliable sellers/providers for future handling, so he doesn't need to remember where he will look for them, figure 15.

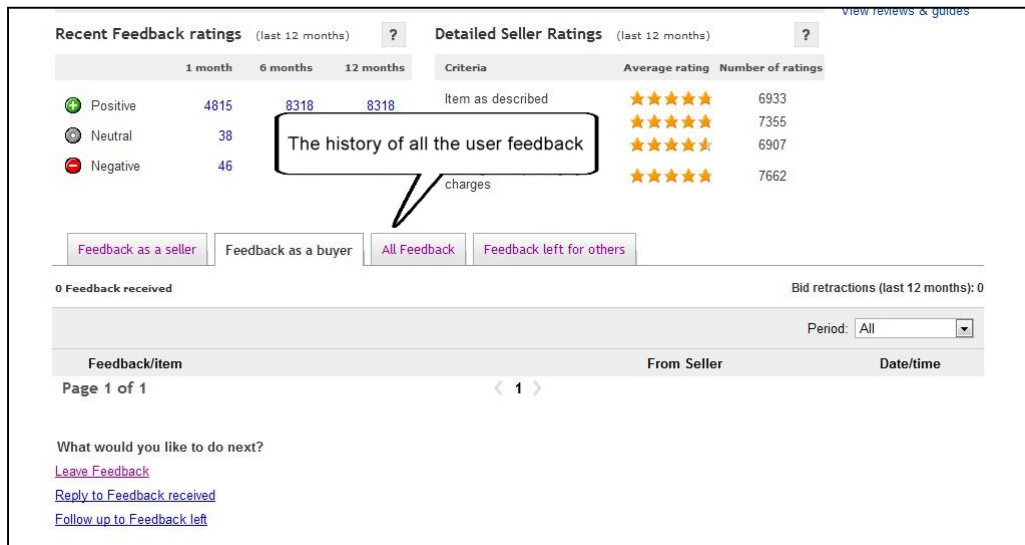


Figure 13: view showing the history of all the member feedback (eBay.com)

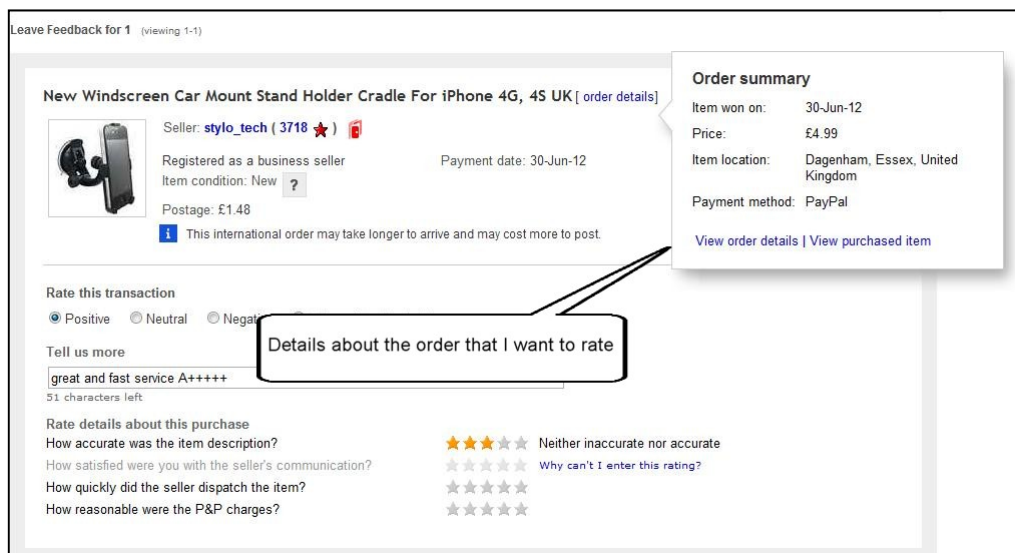


Figure 14: view showing the details of a previous transaction (eBay.com)

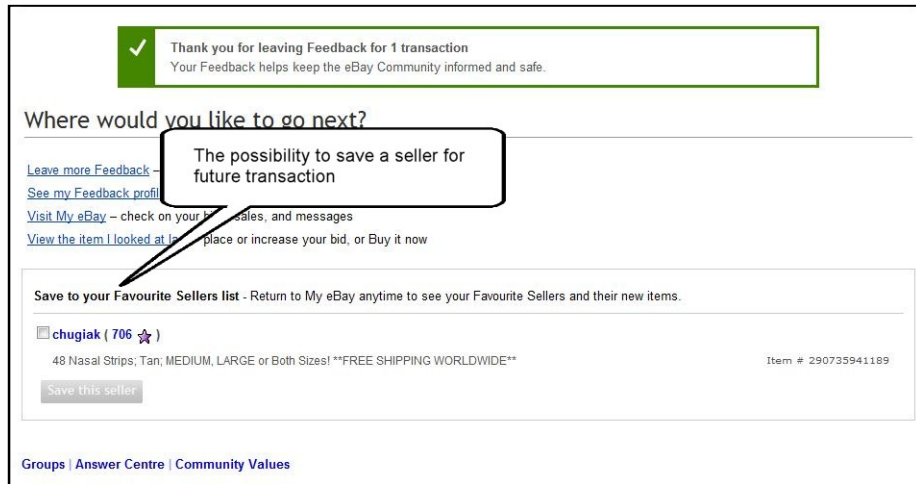


Figure 15: view showing the possibility to save a favorite seller (eBay.com)

## 7. Flexibility and efficiency of use

The system should be flexible to use for every one (novice user, infrequent user and experienced user). The software efficiency is an important aspect of usability. The system should provide some facilities that can have an impact on the overall efficiency of the reputation system and can help users to speed up the interaction with the site, e.g. with just one click at red icon besides visit shop text, the user can visit the seller shop and see what other item he has. Another facility that can speed up the interaction with the site is the Top-rated seller icon; this icon can help the experienced users to identify that the seller has high reputation in the last year, figure 16.



Figure 16: view showing the flexibility and efficiency of use (eBay.com)

## 8. Aesthetic and minimalist design

The system should display all the required (relevant) information at a time. The system should provide related and relevant information about any seller or item, the reputation should be represented according to different aspects, e.g. communication and dispatch time, figure 17 shows how the seller's reputation should be appeared. The user should have the ability to sort the reviews by different categories, e.g. ratings, date and etc. as shown in figure 18. And the navigation should be consistent so the user doesn't feel he is lost, e.g. figure 19 shows the consistent navigation allows the user to know where he would like to go.

The screenshot shows an eBay seller profile for 'locoisle77' with a feedback score of 7829 and a 99.2% positive feedback rate. A 'Detailed Seller Ratings' table is displayed, showing average ratings and the number of ratings for various criteria. A callout box points to the 'Latest Feedback' section, stating 'I can find relevant & enough information about the seller'.

Criteria	Average rating	Number of ratings
Item as described	★★★★★	6932
Communication	★★★★★	7354
Dispatch time	★★★★★	6906
Postage and packaging charges	★★★★★	7661

Figure 17: view showing related and relevant information about the seller (eBay.com)

The screenshot shows a Yelp business page for 'Magasin Du Nord, Kgs. Nytorv'. The page displays business information such as address, phone number, and hours. A callout box points to the review sorting options, stating 'I can sort the reviews by different categories'. The sorting options include 'Yelp Sort', 'Date', 'Rating', 'Useful', 'Funny', 'Cool', 'Total Votes', 'Friends', and 'Elites'.

Figure 18: view showing different categories to sort the reviews (yelp.com)

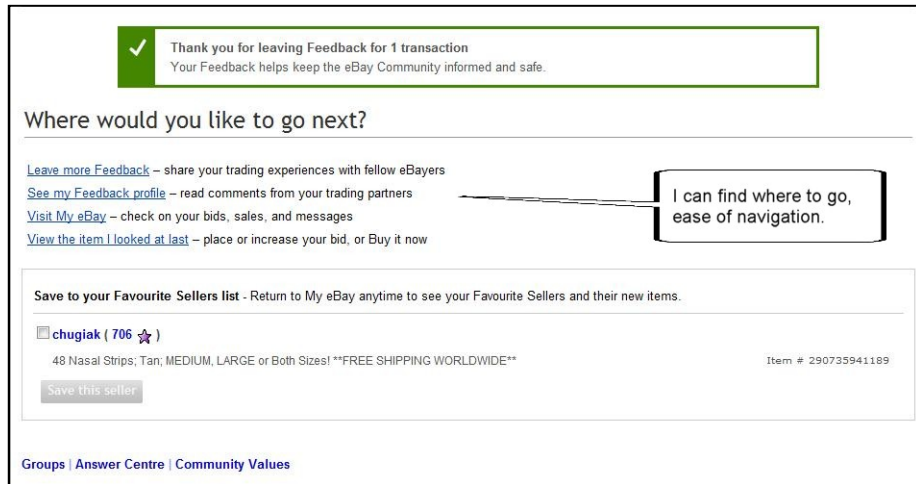


Figure 19: view showing the consistent navigation (eBay.com)

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

The good reputation system should help the user to recognize, diagnose and recover from errors encountered during its use. The help and error messages should be friendly, precise and informative as possible to make it easy for the user to identify the error to recover from it easily. The system should provide help to the users when they need it, and if they want to know about something they don't understand it, there should be help message in plain language that explain the ambiguousness to the user, figure 20 and 21 show how the help messages are in plain language and help the user to understand what he is doing. As in criteria 5 the system should help the user if he does any error by showing an easy plain message.

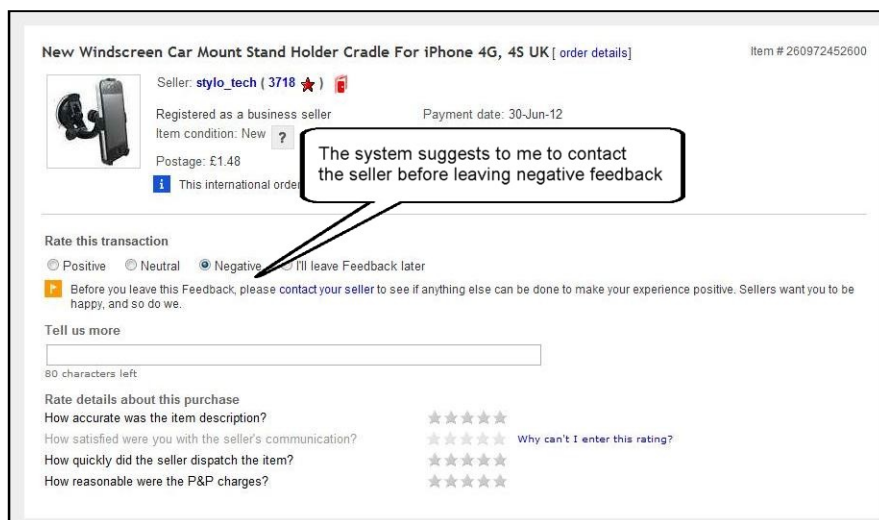


Figure 20: view showing the system's suggestion before leaving negative feedback (eBay.com)



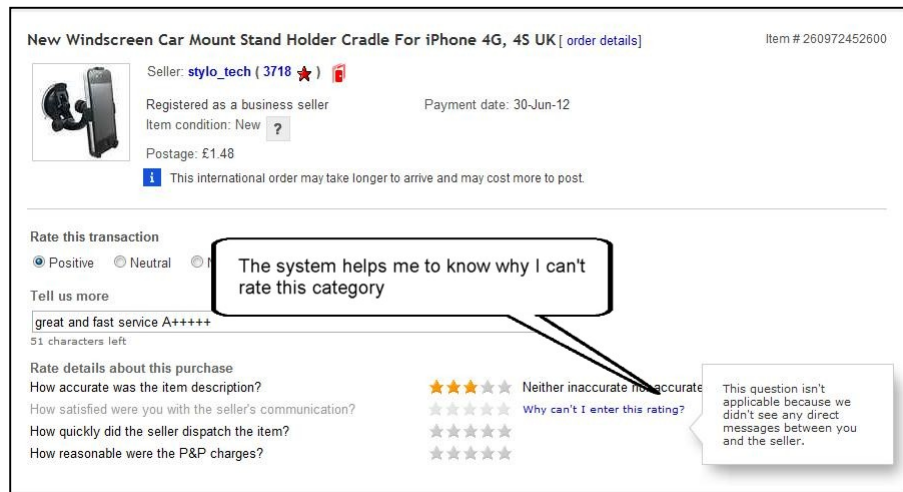


Figure 21: view showing the system explanation to questions (eBay.com)

## 10. Help and documentation

Help and documentation is an important aid for the novice as well as the experienced users. The system should provide help and explanation about anything the users don't understand it. Therefore, if the user is uncertain about any action he should always find the help to resolve the problem. The documentation should fully integrate into a Website. Help could be fully integrated into each page so the users never feel like assistance is too far away. The users should themselves help and figure out their own answers with easy to follow topics or ask the site to help them, e.g. figure 22 shows how the user can find the help in the same page.

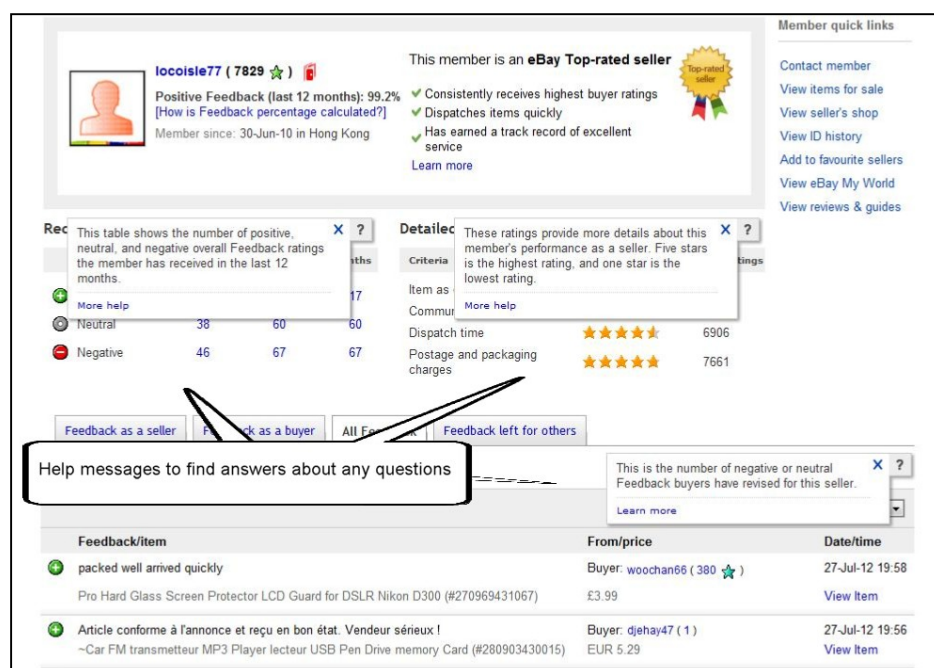


Figure 22: view showing the help messages (eBay.com)

### 3.4.3 Think aloud testing

As the ten heuristics evaluation relies only upon the evaluator's subjective judgment, it would be a better decision to combine it with another usability test. Websites testing with real users is the most fundamental usability method and irreplaceable, since it provides direct information about how people face problems when they use the website. The basis for the usability test and the most valuable usability method is the recognized "think aloud" method. In this testing method test participants are asked to think aloud while they carry out the activity with the system — that is, simply verbalizing their thoughts as they move through the user interface.

The think aloud test consists of 4-20 test sessions. Each test session involves one test participant, who should be a typical user of the website. During the test session, the test facilitator gives the participant the test tasks. While test participants solve their tasks, they are asked to "think aloud" — that is, to say what they are thinking, what they are unsure of, what they expect the website to do, how they interpret error messages, and so on.

Think aloud tests are widely applicable. One can use them to test a beginner's first meeting with a website, as well as to test a long running website with experienced users to see if the site needs updating. One can use think aloud tests on an entire website or on selected pages. One can use them on prototypes and on websites that are in production. A test session normally lasts 60-100 minutes.

This method helps us to discover what users really think about the websites. The main goals of this test are:

- To measure the level of the usability in most common reputation-based sites depending on user point of view.
- To know the most favorable way to present reputation information to the users.
- To examine whether the user understands the rating system and how ratings have been calculated.
- To examine whether the user knows the reputation system rules.
- To examine whether the users are honest enough (don't think of the bad rating's consequence) by leaving a negative/natural feedback.
- To examine if there are more information the user needs in his decision making.

This test has been carried out with six participants who all belong to the target group for the website. The profiles of the test participants appear in section 3.4.3.2.

The test participants were tested one by one. We used remote test with two participants, because they are living outside of Denmark, the other four we tested in our home or their homes.

We acted as the test facilitator. Four tests were conducted in English, while two were conducted in Danish. Each test took between 60-80 minutes. The usability test consisted of three phases: Interview, Solving test tasks, and Debriefing. In this test the phases contained the following steps:

**Interview:** Test participants were informed about the purpose and procedure of the test and were then interviewed about their expectations to the websites before they saw it.

**Solving test tasks:** Test participants were asked to carry out tasks using the websites. All the tasks were defined by us; however some of them made it possible for the test participant to decide about the kind of information they will be looking for. The tasks are included in the usability test script in appendix A. Test participants were asked to think aloud and to comment on the website while they were carrying out their tasks.

**Debriefing:** Test participants were asked to answer some post-test questions in order to summarize their experience with the websites. The list of questions is given in appendix A. [4, 22]

### 3.4.3.1 Equipment

The equipment used for this test were different laptops/desktops with minimum 2.00 MHz processor and a 13.3 – 21.0" wide screen. Google Chrome was used for four test participants, while Mozilla Firefox 3.0.5 was used for the remaining two. There isn't any impact on the test by using two different browsers, the only reason is; the two remote test participants use Firefox, while we use Google chrome. The computer communicated with the Internet using a 10, 20, 35 Mbit xDSL.

### 3.4.3.2 Test Participant Profiles

The tests were carried out with test participants, table 1 who fulfilled the following requirements:

- Are of different nationality.
- Between 25 and 50 years old.
- All of them have experience with browsing the Internet.
- Four of them have not studied computer science or worked within the IT field, the last two studied IT.
- Most of the participants are Danish.

Participant	Gender	Age	Title	Nationality	Internet experience*	Used reputation-based sites
1	Female	26	Physician	American	Experienced	Yes
2	Male	50	Diploma in Applied Chemistry	Danish	Experienced	Yes
3	Female	25	Bachelor in Applied Science	American	Experienced	Yes
4	Female	40	Software Engineering	Danish	Experienced	Yes
5	Male	50	Lawyer	Danish	Experienced	No
6	Female	47	High School Teacher in Physics	Danish	Experienced	No

Table 1: test participants profile

\* Internet experience was classified by the test participant according to these groupings:

1. None (e.g. has ever heard of it or only read about it)
2. Bystander (e.g. has watched other persons use the internet)
3. Beginner (e.g. has used it once or twice)
4. Somewhat experienced (uses it regularly)
5. Experienced (uses search facilities without problems)

### 3.4.3.3 Findings

Findings are categorized by the facilitator using the following categories [4]:



Good. This approach is recommendable.



Good idea. A suggestion from a test participant that could lead to a significant improvement of the user experience.



Minor problem. Caused test participants to hesitate for a few seconds.



Serious problem. Delayed test participants in their use of the website for 1 to 5 minutes, but eventually they were able to continue. Caused occasional "catastrophes".



Critical problem. Caused frequent catastrophes. A catastrophe is a situation where the website "wins" over the test participant, i.e. a situation where the test participant cannot solve a reasonable task or which causes the test participant great irritation.

### 3.4.4 Survey and Interview

In the usability inquiry the usability evaluators obtain information about what users need, like or dislike, so the survey and interviews are very useful methods for studying how users use systems and what features they particularly need, like or dislike. From a usability perspective, survey and interviews are indirect methods, since they do not study the user interface itself but only users' opinions about the user interface. But on the other hand these methods could be considered as direct methods when it comes to measuring user satisfaction. [2]

So we formulate questions about the filtered websites in order to gather information desired. It's the only way to obtain this information by the interactive process between the interviewer and the user. The questions in the survey consist of a modified version of the interview questions.

The main goals of these usability methods are:

- To find what is the important aspect of each user's rating.
- To find whether the user understands the rating system's factors.
- To find whether the user trusts the rating system.
- To examine whether the Ease and Flexibility of the feedback mechanism has an impact on user review.
- To know what the most important information are the users need in their buying decision.
- To examine if there are more information the user needs in his decision making.
- To know what the most things (reputation information representation) that attracts users.
- To examine whether the users are honest enough (don't think of the bad rating's consequence) by leaving a negative/natural feedback.

## 3.5 Reliability and Validity

Reliability and validity are two important aspects in all kinds of testing. The reliability is the question of whether one would get the same result if the test were to be repeated. The validity is the question of whether the result actually reflects the usability issues one to test. [2]

### 3.5.1 **Reliability**

Reliability is concerned with consistency of the results that obtained from the research. The role of the reliability is to reduce the biases in the study. To increase the reliability of our research we have endeavored to make our users sample as representative as possible, and we tried to use open-ended and satisfaction questions.

### 3.5.2 **Validity**

Validity is concerned with using the right users and the right tasks. Validity involves making sure that respondents understand what is meant by the researcher, and all participants should have nearly the identical experience as each other prior to and during the test. In order to increase the validity we have chosen users that know and use the assessed websites. And we have formed clear detail tasks, so the participants can successfully perform these tasks.

# Chapter 4

## 4. Data Presentation

---

This chapter will present the results obtained from the usability studies. The presentation is divided into sections; in the first one we will present the results of the Heuristic Evaluation, in the second section we will present the data obtained from Think aloud and remote testing on the chosen websites and in the last section we will present the results from the survey and the interviews.

### 4.1 Application of Nielsen Ten Heuristics

In this section we will present the findings of our evaluations on the reputation system of the chosen website depending on Nielsen ten heuristics. The findings contain both positives and negatives issues.

#### 4.1.1 eBay.com

eBay is a popular online marketplace; a place for buyers and sellers to come together and trade almost anything. The site allows sellers to list items for sale, and buyers to bid for those items. The eBay reputation system is a centralized reputation system, where eBay collects all the ratings and computes the scores. The eBay reputation system consists of different ratings in the Feedback Profile. The feedback profile has a Feedback Forum that gives buyer and seller the opportunity to rate each other as positive, negative or neutral after completion of a transaction. The buyers and sellers also have the opportunity to rate the seller in 4 additional areas: item as described, communication, shipping time, and shipping and handling charges. These ratings don't count toward the seller's Feedback score, and are anonymous. Detailed seller ratings from the same buyer are counted in the same way as Feedback. Only one per week is included in the seller's score. The buyer has the possibility to leave comments like *"great item quality, will deal again!"* which are typical in positive case or *"Buyers beware, fake item!"* in the negative case.

The most important points in the in the eBay's feedback profile figure 23 are:

1: (Positive Feedback ratings) the percentage of positive ratings left by members in the last 12 months. This is calculated by dividing the number of positive ratings by the total number of ratings (positive + negative).

2: (Recent Feedback ratings) the total number of positive, neutral, and negative Feedback ratings the member has received in the last 1, 6, and 12 months

3: (Detailed Seller Ratings) this rating provides more details about this member's performance as a seller. Five stars is the highest rating, and one star is the lowest. These ratings do not count toward the overall Feedback score and they are anonymous. That means the sellers can't trace detailed seller ratings back to the buyer who left them. Detailed seller ratings from the same buyer are counted in the same way as Feedback. Only one every week is included in the seller's score.

4: (Top-rated seller icon) this icon means that the seller is:

- Consistently receives highest buyer ratings.
- Dispatches items quickly.
- Has earned a track record of excellent service.

5: (Bid retractions) the number of times the member has retracted a bid in the last 12 months. To see the number of bid retractions, click the "Feedback as a buyer" tab.

6: (Textual Feedback) this refers to the comments that are left from other members.

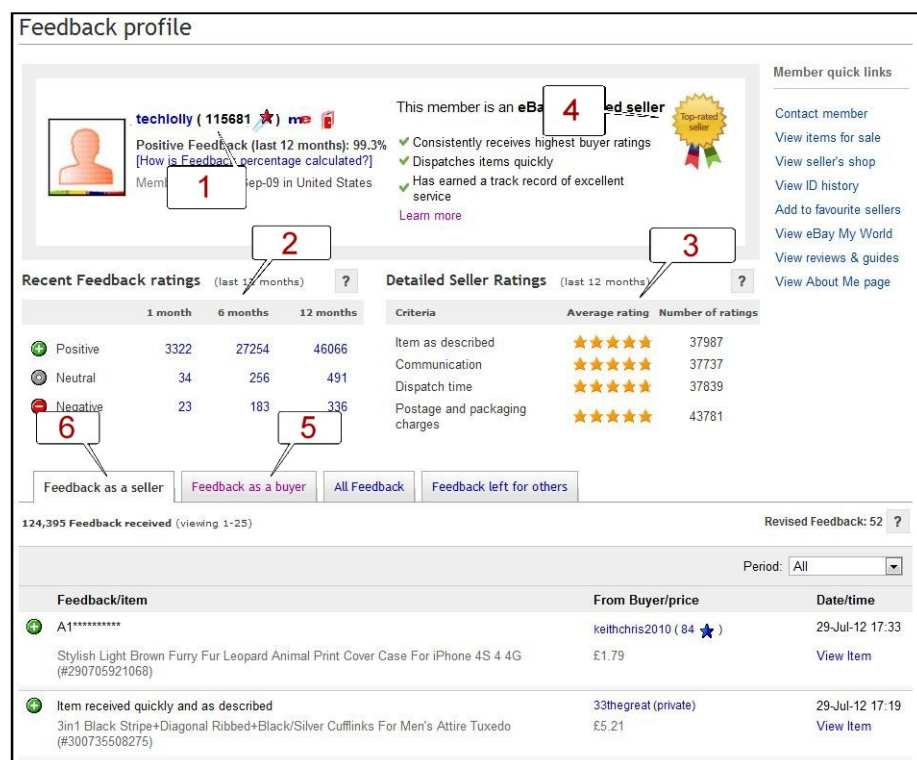


Figure 23: the eBay feedback system

The problem of ballot stuffing, i.e. that ratings can be repeated many times, e.g. to unfairly boost somebody's reputation score, seems to be a minor problem on eBay because participants are only allowed to rate each other after the completion of a



transaction, which is monitored by eBay. It is of course possible to create fake transactions, but because eBay charges a fee for listing items, there is a cost associated with this practice. However, unfair ratings for genuine transactions cannot be avoided. [5, 6]

## Nielsen heuristics for eBay's Reputation Systems

There are some negative and positive points in the eBay reputation system depending on our Nielsen heuristics evaluation. The most important points are; the user always knows where he/she is in terms of carrying out a procedure using a system, Users would know if an operation was successfully completed (e.g., you left a feedback), figure 6. eBay reputation system uses natural words, particularly novice users can understand the language and the system categories are well grouped as shown in figures 7 and 8. One of the negatives in the eBay reputation system is; there is no undo functions in this system, if the user leaves a feedback and later he regrets it, he cannot change it. The users can find all the information they need quickly and they can understand what every rating means, e.g. they can understand how the feedback percentage is calculated by clicking at the link under the positive feedback as shown in figures 10 and 11. The eBay system can prevent an error but only for PowerSellers\*, so if a member by mistake wants to leave a negative feedback to a power seller the system will prevent him from doing this before 7 days (e.g. Sorry, you can't leave negative or neutral Feedback for this PowerSeller until 7 days after purchase. in the mean time, please contact the seller to try to work things out.) as shown in figure 24.



Figure 24: eBay prevents users from leaving a negative feedback to powerseller

In the eBay system the users can retrieve all the history of other users' feedback, and they can see all the details about any transaction before they can rate it as shown in figures 13 and 14.

\*PowerSeller: he is an eBay seller who sustaining a gross trading volume above a set cutoff for several months in a row, Power sellers must maintain both a high quality feedback profile and constant or growing trading volume in order to remain in the program.

The eBay system provides some facilities that help users to speed up their interaction with the system such as “visit shop” as shown in figure 16. The eBay system provides also relevant information about any seller figure 17, but unfortunately the users can only sort other users' reviews by “date”, there is no other category. The eBay system helps users to find out what additional steps are needed in order to complete a task successfully, for example the question “How satisfied were you with the seller's communication” you can't rate it sometimes, and if you want to know why that? The answer directly comes to you: (This question isn't applicable because we didn't see any direct messages between you and the seller) as shown in figure 21. The eBay system has a very good help, it provides a quick and accurate help as shown in figure 22.

#### 4.1.2 Amazon.com

Amazon is a popular online shopping for everything, although it starts mainly as an online bookstore that allows members to write book reviews. Anybody can become a member simply by signing up. The seller rating system is 1 to 5 stars, with 5 stars being the best. A seller's average rating will appear alongside his name. Buyers have 90 days after their order date to leave their rating and remarks. The average of all ratings gives a seller or a product its average rating. Users, including non-members, can vote on product's reviews as being *helpful* or *not helpful*. The numbers of helpful as well as the total number of votes are displayed with each review. The order in which the reviews are listed can be chosen by the user according to criteria such as newest first or most helpful first or depending on the number of the stars rating. While the seller reviews can only be sorted by date and can't be sorted by any other criteria.

Amazon has the right to remove feedback that is not directly related to the buying experience or violates one of Amazon guidelines. If the comments include any of the following, the feedback is subject to removal:

- **Product reviews:** It is more appropriate to review product on the product detail page.
- **Promotional content:** This includes anything of a promotional nature such as comments about or links to other merchants or websites.
- **Obscene or abusive language:** this mean the member should use helpful and appropriate language when participating in the Amazon Community.
- **Personal information:** members should not include information that identifies other Amazon.com visitors.

Amazon can offer reputation to stores that reside on its site. Amazon is a trusted merchandiser, and consumers assume that any site Amazon allows on its web pages has been vetted. This halo provides third-party credibility to merchants who might not be

able to gain good reputations alone in a short time. Merchants, of course, pay for the privilege of borrowing a good name.

The most important points in the in the Amazon's reputation profile figure 25 are:

- 1: (Star Ratings) the seller stars is an average rating calculated using the following formula:  $\text{SUM}(\text{all feedback star values}) / \text{SUM}(\text{all feedback ratings})$ . For averages over 4.76, all 5 stars are shaded; 4.26 to 4.75 = 4 1/2 stars, 3.75 to 4.25 = 4 stars, and so on
- 2: (Detailed Seller Information) details information about the seller, e.g. dispatch time and offering warranty to his products.
- 3: (Recent Feedback Ratings) the recent reviews are illustrated by date with its rating.
- 4: (Feedback History) the feedback percentage is calculated by using the following methodology:  $\text{SUM}(\text{positive feedback ratings}) / \text{SUM}(\text{all feedback ratings})$  for feedback left in the last 30, 90, 365 days, and lifetime. [1, 7]

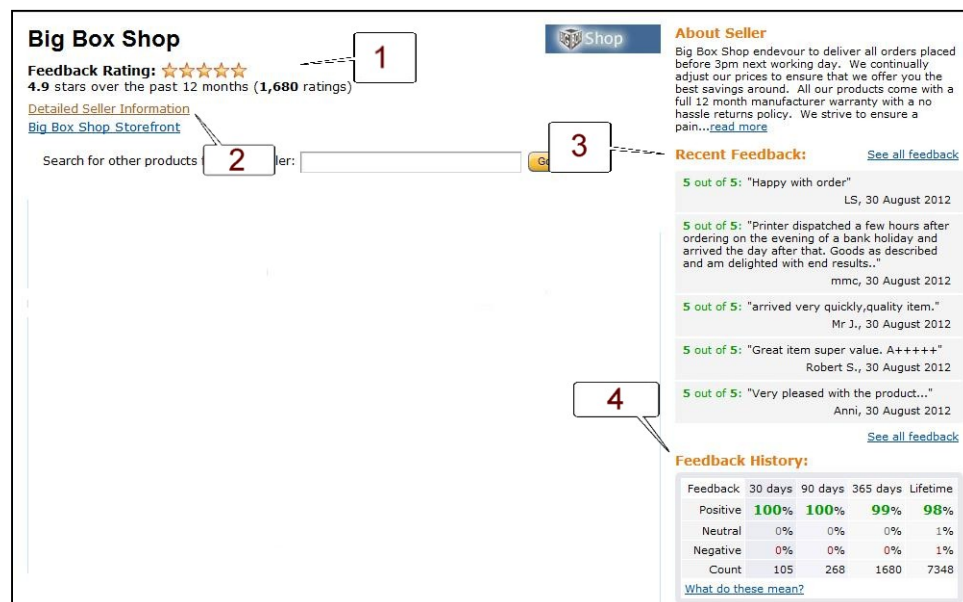


Figure 25: the Amazon reputation system

## Nielsen heuristics for Amazon's Reputation Systems

There are some negative and positive points in the Amazon reputation system depending on our Nielsen heuristics evaluation. The most important points are; the user always knows where he is and what he is doing in terms of carrying out a procedure using a system figure 26, Users would know if an operation was successfully completed (e.g., Thanks! Your review is being processed) figure 27.

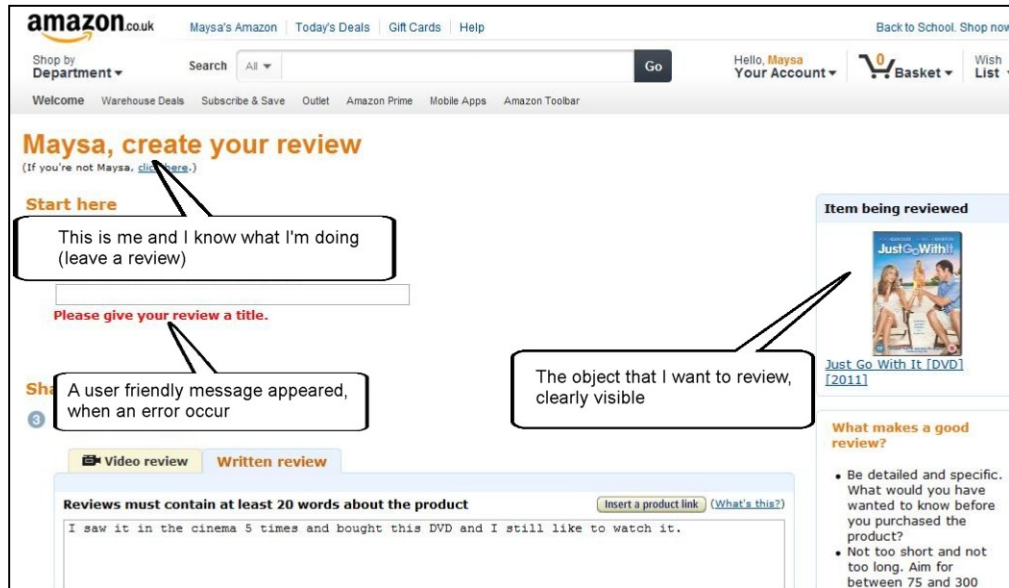


Figure 26: view showing account name and the reviewed object plus a friendly message (Amazon)

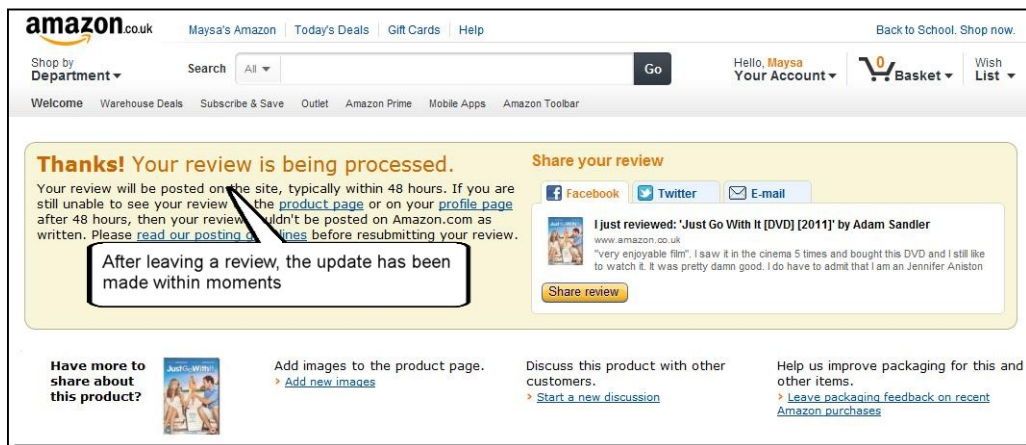


Figure 27: Amazon provide the updates within moments

Amazon reputation system uses natural words, particularly novice users can understand the language and the system categories are well grouped. One of the negatives in the Amazon reputation system is; there is no redo/undo mechanism in the Amazon reputation system, so you cannot change or comment your earlier feedback, although you can remove the negative ones but once you remove it you can't post a new feedback instead of the removed one as shown in figure 28.

Your Submitted Feedback					
Order Number	Seller Name	Rating	Comments	Date	
203-2151623-8154714	zoverstocks	3 out of 5	"the seller was great even if the dvd didnt work with me (vista windows), so the seller refund me the price of the dvd."	12 July 2012	Remove
202-6513626-8973143	World of Books Ltd	5 out of 5	"the seller very honest and he sent me the book quickly. great service. and good book."		
202-1437117-1944338	Nearfine_	5 out of 5	"He is a perfect seller"		
026-5724425-8702735	Pharmacy Place	5 out of 5	"great service"	29 January 2012	
026-6040726-8657119	Beauty Base Online	5 out of 5	"great service"	29 January 2012	

I can remove only the natural and negative review

Figure 28: Amazon provides only remove choice to negative reviews

The Star rating system in Amazon is not very clear (e.g. when a user rates a book with bad, users might think it was badly written, while in reality the user might not like the topic for example), and the information of the reputation system is presented similarly, so the user sometimes gets confused where the item details are and where the users review and other information are. This makes retrieving information for the decision-making a hard task as shown in figure 29.

### Product Description

Review

**'Usability Engineering** is a good introductory text that is readable and thought provoking...It promotes active, independent learning by encouraging students to explore and question concepts related to usability. I recommend this book as a starting point for students, or IT managers alike, who wish to get an insight into the practicalities of systems usability.' - Kathleen Maitland, Lecturer in Computing, UCE at Birmingham

About the Author

XRIS FAULKNER is a Senior Lecturer at South Bank University where she has taught for the past seven years.

---

### Inside This Book [\(Learn More\)](#)

**First Sentence**  
One of the biggest problems that the designers of a human computer system face when developing a system is making sure that the finished product is what the user really wants and needs. [Read the first page](#)

**Explore More**  
[Concordance](#)

**Browse Sample Pages**  
[Front Cover](#) | [Copyright](#) | [Table of Contents](#) | [Excerpt](#) | [Index](#) | [Back Cover](#)

**Search inside this book:**

---

**Tag this product** [\(What's this?\)](#)

Think of a tag as a keyword or label you consider is strongly related to this product. Tags will help all customers organise and find favourite items.

Your tags:

(Press the "T" key twice to quickly access the "Tag this product" window.)

**Search Products Tagged with**

[> See most popular tags](#)


---

**Sell a Digital Version of This Book in the Kindle Store**

If you are a publisher or author and hold the digital rights to a book, you can sell a digital version of it in our Kindle Store. [Learn more](#)

---

**What Other Items Do Customers Buy After Viewing This Item?**



Usability Engineering (Interactive Technologies) by Jakob Nielsen Paperback

★★★★☆ (5)

£17.00

Figure 29: the reputation information is presented similarly in Amazon

In the Amazon system the users can retrieve all the history of other users' feedback, but one can't for example choose to retrieve only the negative ones, so one might navigate 100 pages only to reach a negative review figure 30.



Figure 30: no review's classification in Amazon

The Amazon system provides some facilities that help users to speed up their interaction with the system such as “customer reviews” as shown in figure 31.



Figure 31: Amazon provides some facilities that speed up the interaction with the system

The Amazon system doesn't provide enough information about any seller, and it's difficult to navigate people profiles, and the users can only sort other users' reviews by “date”, there is no other category, but the good thing in the Amazon review system; it provides detailed rating's information about how many users rated the item and gives a direct access to their reviews depending on different criteria such as the most helpful first review or the newest first figure 32.

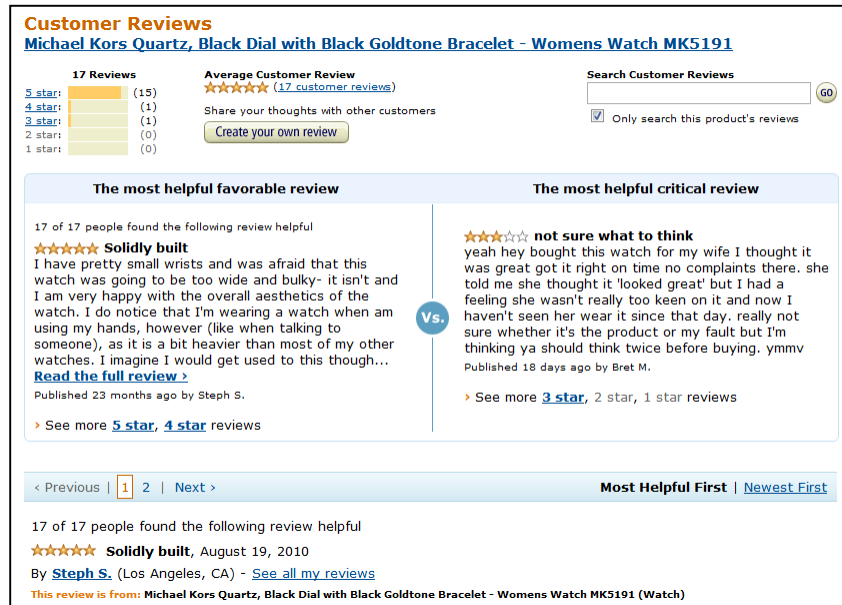


Figure 32: Amazon provides sorting item reviews by different criteria

The Amazon system help users to find out what additional steps are needed in order to complete a task successfully, and if an error does occur, the system provides a user friendly message in plain language, e.g. if the user wants to leave a review and he forgets to enter a title for the review, a user friendly message will appear; “please give your review a title” figure 26. Amazon has a very good help system, it provides an accurate help. The user can use self help and figure out his own answers with the easy to find and press on the link in any place figure 33.

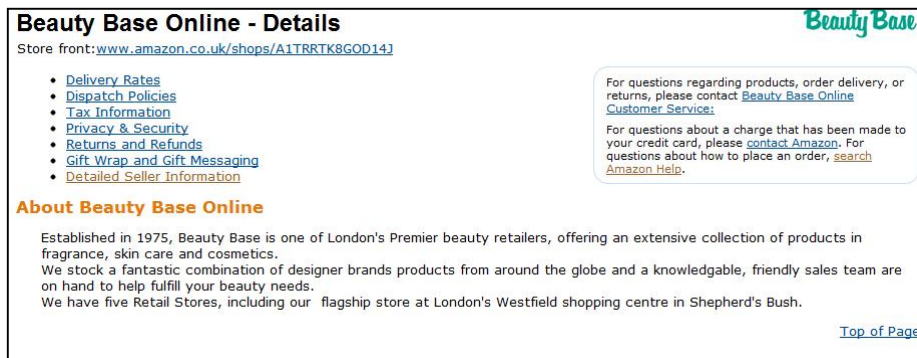


Figure 33: Amazon provides a good help system

### 4.1.3 Yelp.com

Yelp is an online urban city guide that helps its visitors find local cool places to eat, shop, drink, relax and play, each place has a 5-point rating, reviews from other site visitors. Yelp helps people to find, review and talk about what's great — and not so great — in our world. Any local business, service or place with a physical presence such as restaurants, shops, bars, salons, spas, dentists, mechanics, parks, museums, etc. The Yelp Elite Squad is Yelp way of recognizing and rewarding yelpers who are active evangelists and role models, both on and off the site. Elite-worthiness is based on a

number of things, including well-written reviews, a fleshed-out personal profile, an active voting and complimenting record, and playing nice with others. Members of the Elite Squad are designated by a shiny Elite badge on their account profile, so the visitors can see which members are in Yelp Elite Squad.

Yelp provides a feature that allows business owners to publicly comment on a review. Members are able to message them back if they post a public comment on their review.

Users can vote on reviews that they think are Useful, Funny or Cool. Users can also decide for themselves how best to order reviews by clicking one of the links just above the reviews (e.g., date, rating, voting, etc.). Yelp's default sort order takes a number of factors into account and present reviews in a meaningful order. For example, Yelp will favor reviews from the member friends and the users he/she follows. The sort algorithm does not take into account whether the business is an advertiser or not.

Sometimes yelp may remove a review if it violates their Terms of Service, and a reviewer may remove her own review of her own accord. Yelp has "Filtered Review" mechanism trying to showcase the most helpful and reliable reviews among the millions that are submitted to the site. Filtered reviews don't factor into a business's overall star rating, but users can still read them by clicking on the link at the bottom of the business's profile page.

The most important points in the in the Yelp's rating profile figure 34 are:

1: (Star Ratings) the rating is calculated by using Bayesian approach, we talked about it in section 2.3.

2: (Detailed Business Information) details information about the business, e.g. address, opening hours, price range.

3: (Review Highlights) these phrases are mentioned a lot by users. These phrases are also the Yelp Robots have determined are unique and good, quick ways to describe this business.

4: (Rating Details) there are two graphs. First one is the distribution of the 1, 2, 3, 4 and 5-star ratings. The second graph is the average rating and how it has changed over time.

5: (Textual Reviews) this refers to the comments that are left from other members, and can be sorted by different categories, e.g. date and rating. [8]



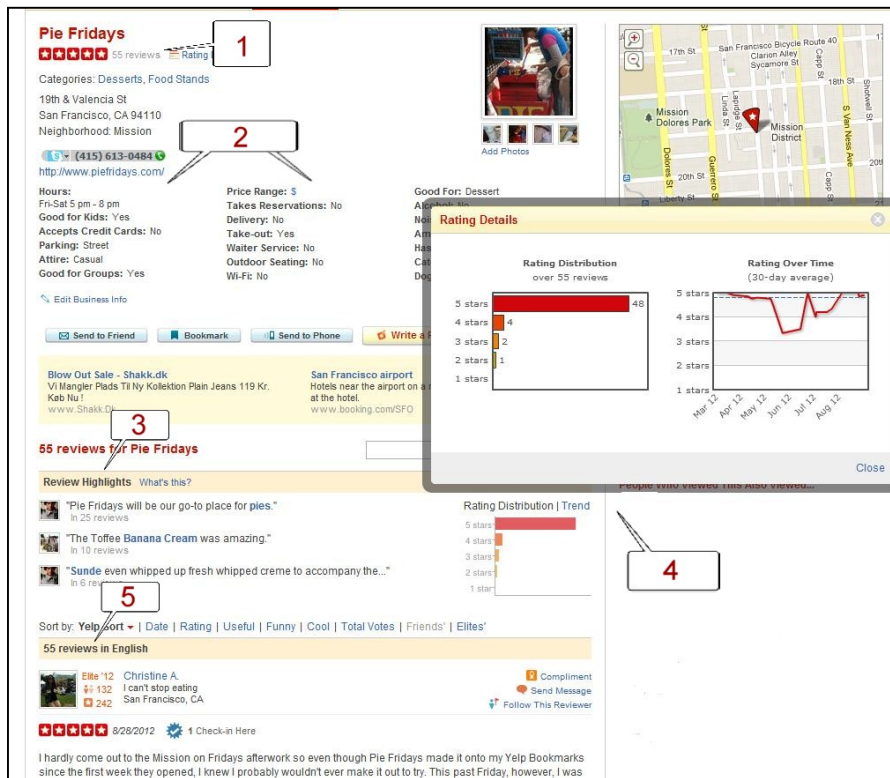


Figure 34: Yelp rating information

## Nielsen heuristics for Yelp’s Reputation Systems

There are some negative and positive points in the Yelp reputation system depending on our Nielsen heuristics evaluation. The most important points are; the user always knows where he is and what he is doing in terms of carrying out a procedure using a system figure 5, Users would know if an operation was successfully completed e.g. after rating a review the number of ratings is immediately updated, which shows the reliability of the system, and when you leave a review, a message appears within moments “Your review for (the business) has been posted” figure 35.



Figure 35: provide the updates within moments

Yelp reputation system uses natural words; particularly novice users can understand the language. One of the positives in the yelp reputation system is; the system provides the option to undo (edit or remove) the ratings and the reviews the user did before to the business; this will allow the user to correct any mistakes have been made and will add more trustworthiness to the ratings as shown in figure 9.

Yelp provides a graphical and clear distribution of ratings, but it doesn't show how the rating score is obtained. The review contains much information about the reviewers' reputation, making the relevant information available to the user and the reputation of the review itself can also be seen, but there is some negatives such as the rating system does not provide clear information about the quality of the place (e.g. the user can see many reviews of a restaurant but cannot see its quality compared with another restaurant in the same city).

In the yelp system the users can retrieve all the history of other users' reviews, and the review contains much information about the reviewer reputation, so the members can see how the review and the reputation of the reviewer are reliable.

The yelp system provides some facilities that help users to speed up their interaction with the system such as "Compliment", "Send Message" and "Follow This Reviewer" as shown in figure 36.

The yelp system provides relevant information about any business, and the user can sort other users' reviews by many categories e.g. date, rating and useful and etc., but there are some negatives with the sort. The Yelp Sort is not clear; which doesn't clearly show what a sort entails. When press at "Yelp Sort" it shows a menu explains the default sorting schema, while the user expects multiple-choice option. So a new user might see the reviews in no particular order. Another problem with rating sort category; the arrows beside the rating category are not clear whether they refer to negative or positive reviews figure 36.

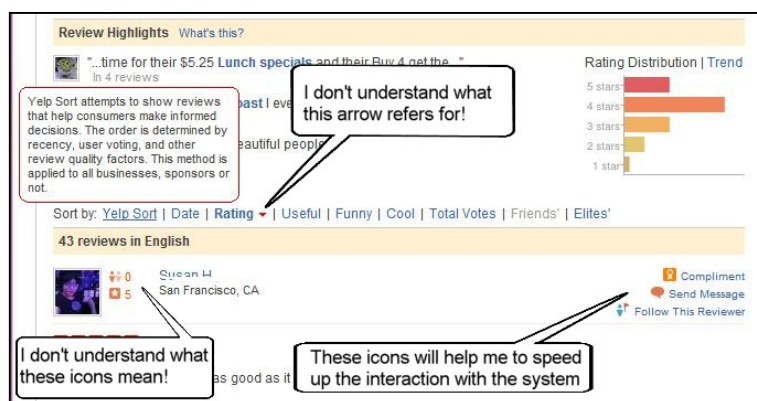


Figure 36: yelp provides some facilities but has some negatives

The yelp system helps users to find out what additional steps are needed in order to complete a task successfully, and if errors do occur, the system provides a user friendly message in plain language, e.g. if the user wants to leave a review and he forgets one of

the fields, a user friendly message will appear; "A required field was not filled." As shown in figure 37.



Figure 37: yelp help users to recover from errors

Yelp has a moderate help system, in some places it provides an accurate help. The user can use self-help and figure out his own answers with the easy to find and press on the link in many places, and sometimes the user cannot understand what an icon means figure 36.

#### 4.1.4 Trustpilot.com

Trustpilot is a portal gathering reviews on online shops. It is an open, review- driven community connecting online consumers with the companies that the consumer buy from. Customer reviews are the core of the Trustpilot concept. Every review helps consumers make more informed and smarter decisions when shopping online. All reviews are open to the public. The reviews can be edited if it is based exclusively on the product features or the layout/usability of the website, or there is documentation that parts of the review are deliberate misrepresentations of the actual course of events, or it contains names of employees, coarse language or references to other companies (such as competitors, etc).

The reviews can be removed if the reviewer has not ordered a product or service at the company, or has not had a user profile at, for instance, dating sites or it has been written by the owner or employees of the company, or the reviewer has written more than one review about the company, or posted duplicate reviews. But the reviews cannot be removed if the company and the reviewer disagree about the review, and where it's impossible to determine who is right. In the like of this case, the company must reply to the review and state their side of the matter.

The TrustScore is an average that takes representatively and the reviews' "age" into account. It's not simply an average of individual user ratings. It attempts at giving a

mathematically sound approximation of the probability of having a satisfying experience when shopping with a given company. One thing it does to achieve this, is to never allow a rating to deviate too far from the average, unless they have enough data to soundly make bold claims about the customer satisfaction. If there are only few reviews, the score will not deviate a lot from the median, 7. As more reviews are written, the formula functions more and more like a weighted average, where new reviews count more than old reviews.

There are many users are verified "Verified buyer", which means that TrustPilot have been in contact with the user who has verified his/her purchase with the shop in question. Usually, they do this by forwarding a scanned copy of the order confirmation or invoice, or by forwarding their email order confirmation. Furthermore, companies who have signed up as partners have the option of sending feedback emails to their customers. When customers write reviews by clicking the link in these emails, they will automatically be verified buyers. [9]

The most important points in the in the Trustpilot's reputation profile as shown in figures 38 and 39 are:

- 1: (Star Ratings) TrustScore is very complicated and includes several parameters. The TrustScore is a means to sum up consumers' shopping experiences with any given company into a single and comparable score ranging from 0 to 10. It's calculated as a weighted average of all registered reviews of a company with multiple parameters taken into account including: Number of reviews, age of the review, aggregated reviews from other sites, trustworthiness of the user writing the review and more.
- 2: (Rating distribution): it shows the distribution of rating for the business from different portal gathering reviews websites.
- 3: (Recent Review) the recent reviews are illustrated by date with its rating.
- 4: (Detailed business Information) details information about the business.
- 5: (Business ranking) it shows the business ranking in the world in their field depending on the trust score from the users.

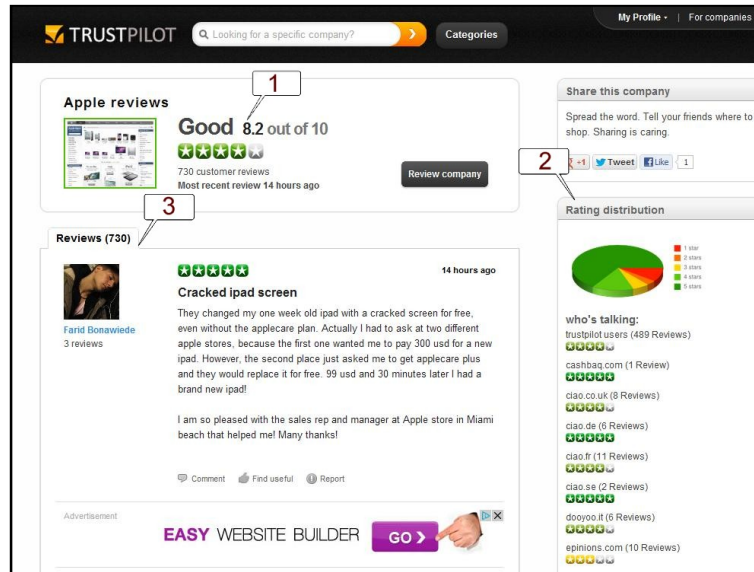


Figure 38: TrustPilot Rating system (part 1)

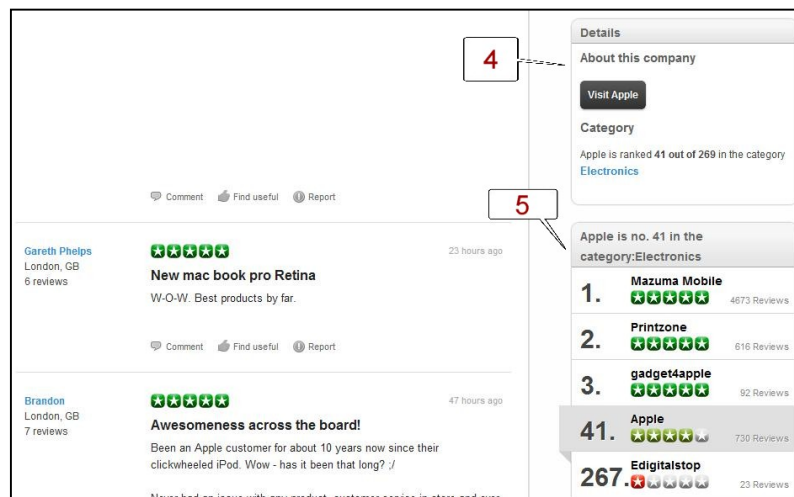


Figure 39: TrustPilot Rating system (part 2)

## Nielsen heuristics for Trustpilot's Reputation Systems

There are some negative and positive points in the TrustPilot reputation system depending on our Nielsen heuristics evaluation. The most important points are; the first negative in trustpilot system is the user cannot know where he is, e.g. user's name doesn't appear when he wants to leave a review, but the user does know what he is doing in terms of carrying out a procedure using a system figure 40.

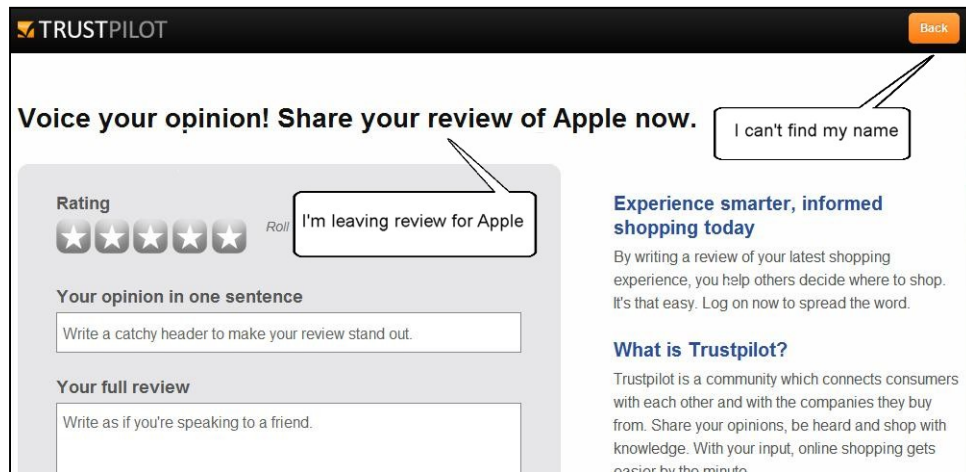


Figure 40: Trustpilot system visibility

Users would know if an operation was successfully completed (e.g., “username”, Thanks for your review of “Business name”) figure 41.

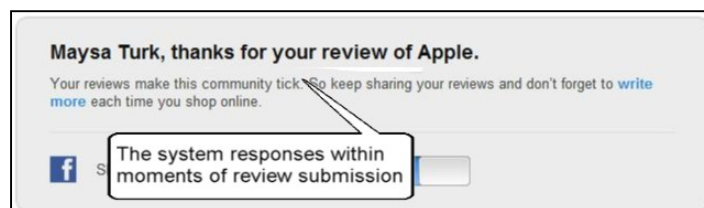


Figure 41: Trustpilot provides updates within moments

Trustpilot reputation system uses natural words; particularly novice users can understand the language. One of the positives in the Trustpilot reputation system is; the system provides the option to edit or delete the reviews the user left earlier. This will allow the user the possibility to correct any mistakes that have been made before or maybe he changes his mind about the reviewed company, this will add more trustworthiness to the ratings as shown in figure 42.

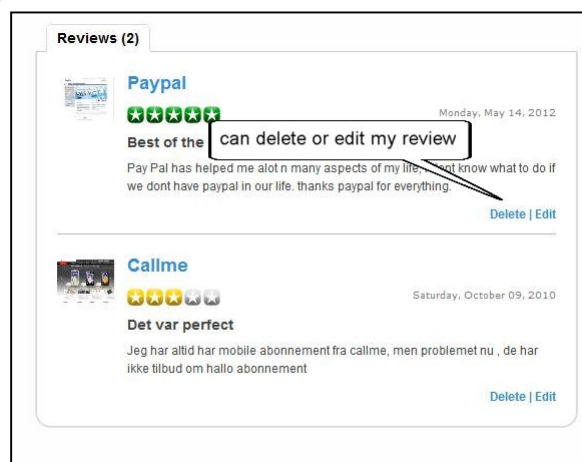


Figure 42: the possibility to change reviews in Trustpilot

The Star rating system in Trustpilot has some negatives such as the user doesn't know how the rating score is calculated, but on the other hand there are some positives such as the system provides rating distribution and business ranking about the business so the user can know more about the business. One of the positives in Trustpilot; for negative comment you can find the answer form the reviewed company. Another positive point the Trustpilot system explains clearly what each star means as shown in figure 43.



Figure 43: clear understanding to the star rating in Trustpilot

In the Trustpilot system the users can retrieve all the history of reviews, and the review contains information about the reviewer reputation e.g. verified buyer, and also there are a possibility to rate a review with a useful or to report it. This will add trustworthy to the reviews figure 44.

The Trustpilot system provides some facilities that help users to speed up their interaction with the system such as "Comment"; to leave a comment about a review as shown in figure 44.

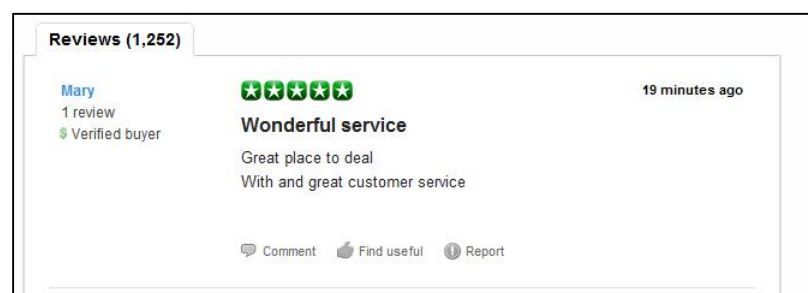


Figure 44: Trustpilot provides some facilities

Trustpilot provides enough information about the business, but unfortunately the users can only sort the reviews by "date", there is no other category.

The Trustpilot system helps users to find out what additional steps are needed in order to complete a task successfully, and if errors do occur, the system provides a user friendly message in plain language, e.g. if the user wants to leave a review and he forgets to enter a headline for the review, a user friendly message will appear; "please write a headline with more than 3 characters." figure 12.

Trustpilot has a moderate help system, in some places it provides an accurate help, e.g. by putting the cursor on the ambiguous icon and sometimes you can't find a help so you need to search about it.

#### 4.1.5 **Epinions.com**

Epinions is one of the most famous product and shop review sites. This site has a pool of individual reviewers who provide information for consumers for the purpose of making better purchase decisions. The reputation system in Epinions applies to products as well as to the reviewers themselves. Epinions is based on so-called cost-per-click and Pay-Per-Click online marketing, which means that Epinions charges product manufacturers and online shops by the number of clicks consumers generate as a result of reading about their products on Epinions web site. Also Epinions pay the reviewer each time someone clicks onto reviewer article, each visitor only registers once, so even if they read it again on a different day, you will only be paid the first time the reader visits the article.

Epinions has a pool of members who write product and shop reviews. Anybody from the public can become a member simply by signing up. The product and shop reviews written by members consist of prose text and quantitative ratings from 1 to 5 stars. The review must be a minimum of 20 words and pass Epinions' automated language check, and the member may only write one review per product or service, but he can add to his review on the product, however, he may update his review at any time, and the rating consists of a set of aspects such as *Ease of Use*, *Battery Life* etc. in case of products, and *Ease of Ordering*, *Customer Service*, *On-Time Delivery* and *Selection* in case of shops.

Other members can rate reviews as *Not Helpful*, *Somewhat Helpful*, *Helpful*, and *Very Helpful*, and members can change their rating. The rating doesn't count equally for all members; it means Advisors have a greater impact on the overall ratings. A member can obtain different status:

**Advisor:** they are active members of the Epinions community who help shoppers find the best content on Epinions by rating reviews in their category. Advisors also provide constructive feedback via comments to reviewers on how to improve content quality, and they carefully choose whom to trust.

**Top Reviewer:** are active members of the Epinions community who help shoppers find the best products on Epinions by writing high quality reviews in their category of expertise. Top Reviewers' reviews have received the highest ratings from the Epinions community, and they carefully choose whom to trust.

**Category Lead (highest):** are active members of the Epinions community who help Epinions oversee a particular category. There are three tasks which Category Leads are required to fulfill, they are:



- Ensure that new reviews in their category are rated by a Category Lead or a category Advisor.
- Work to increase high-quality review coverage of key products in their category.
- Assist with selection of Top Reviewers and Advisors for their category.

Category Leads also perform other functions, but these are not required of them. These include, but are not limited to:

- Adding products which are not listed on Epinions and other members wish to review.
- Recruiting and mentoring new members.
- Acting as the point of contact for Epinions to discuss member concerns.
- Preview and offer feedback on new developments at Epinions.com.

It takes considerable reviewing effort to obtain a status above member, and most members don't have any status.

Category Leads are selected at the discretion of Epinions staff each quarter based on nominations from members. Top Reviewers are automatically selected every month based on how well their reviews are rated. Advisors are selected in the same way as Top Reviewers, but with a lower threshold for review ratings. Epinions does not publish the exact thresholds for becoming Top Reviewer or Advisor, in order to discourage members from trying to manipulate the selection process.

The most important points in the in the Epinions's reputation profile figure 45 are:

1: (Star Ratings) the star rating in Epinions is calculated based on the following criteria:

- Average product rating, with extra weight given to high quality reviews.
- Number of reviews about the product.
- Recency of reviews about the product.

2: (Place and Price) details information about the shops that sell the chosen product with comparison of its prices.

3: (Textual Reviews) this refers to the reviews that are left from other members, and can be sorted by date or rating.

4: (Detailed Product Information) details information about the product. [1, 10]

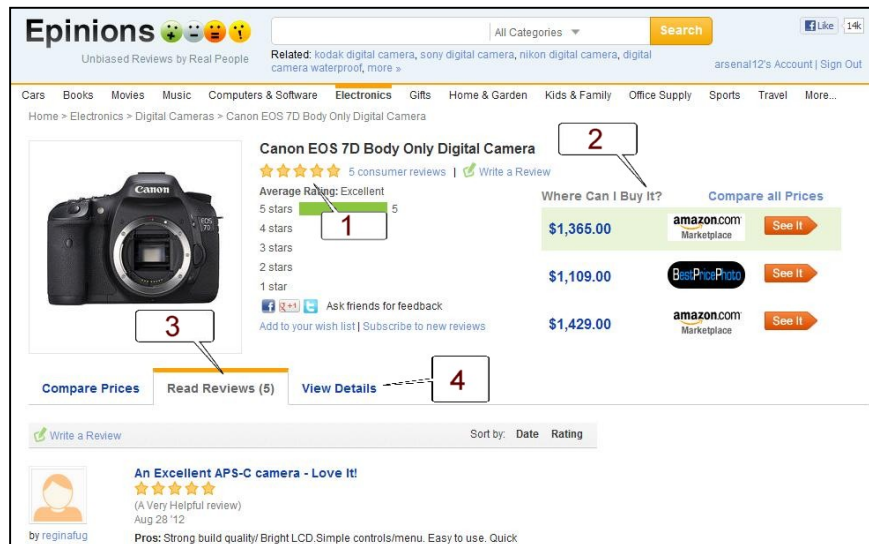


Figure 45: Epinions rating system

## Nielsen heuristics for Epinions's Reputation Systems

There are some negative and positive points in the Epinions reputation system depending on our Nielsen heuristics evaluation. The most important points are; the user always knows where he is and what he is doing in terms of carrying out a procedure using a system, e.g. when a user writes a review on a product as shown in figure 46.

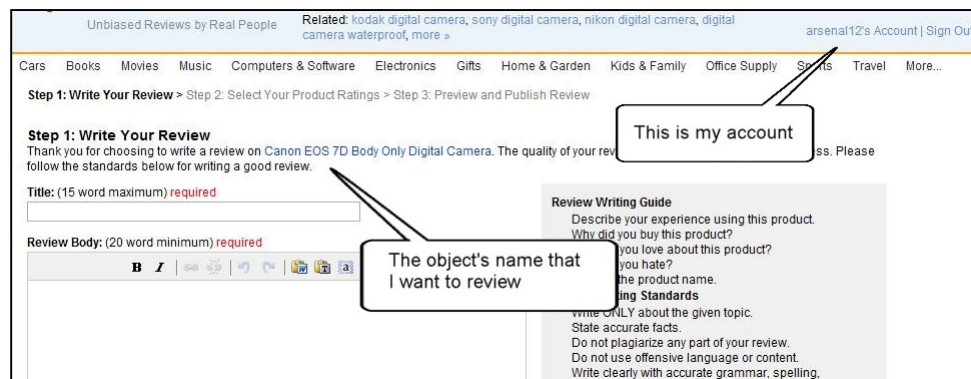


Figure 46: Epinions system visibility

Users would know if an operation was successfully completed e.g. after rating a review a message will appear that shows which rate the users choose as shown in figure 47, another example when user writes a review, a message will appear within moments “Thank you. Your review has been submitted and published” as shown in figure 48.



Figure 47: Epinions update within moments



Figure 48: Epinions responds of review submission

Epinions reputation system uses natural words; particularly novice users can understand the language. One of the positives in the Epinions reputation system is; the system provides the option to undo (update, move or delete) the ratings and the reviews the user did before to the product; this allow the user to modify his review if he changes his mind and this will add more trustworthiness to the ratings as shown in figure 49.

The star rating in Epinions is not understandable, and user cannot figure out how the score has been calculated, and for what it refers to, but on the other hand when users read authors reviews they can see the rating of each category and what the rating refers to such as Durability or Clarity as shown in figure 49.

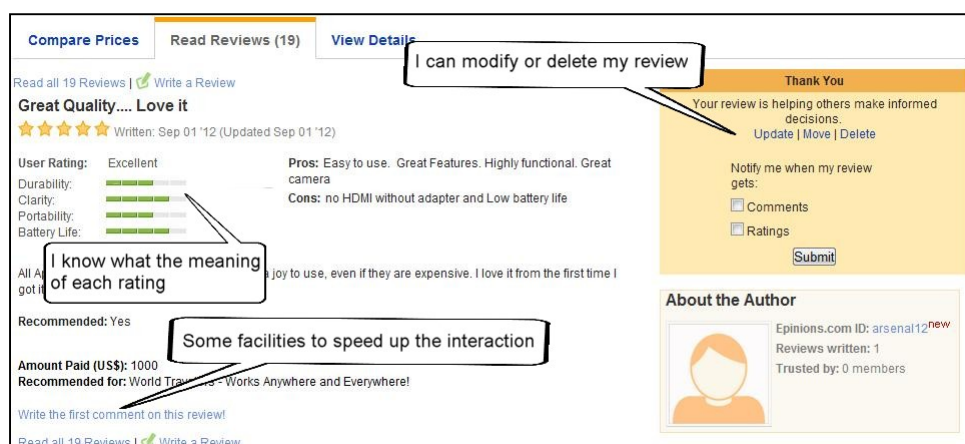


Figure 49: Epinions system provides user control

The review contains much information about the author's reputation, making the relevant information available to the user and the reputation of the review itself can also be seen e.g. A very Helpful review. The other positive point in Epinions; the system

provides obvious information about the quality of the product and gives a comparison of its price in different shops.

In the Epinions system the users can retrieve all the history of other users' reviews. The Epinions system provides many facilities that help users to speed up their interaction with the system such as "Read all Reviews", "write a Review" and "write comment on this review" as shown in figure 49.

The Epinions system provides relevant information about any product and the user can sort the reviews according to "date" or "rating", and the review contains much information about the reviewer reputation, so the members can see how the review and the reputation of the author are reliable as shown in figure 50 , but there is a negative point when a user wants to write a review, it requires from him to write many fields, title with maximum 15 words, review body with 20 word minimum, pros with 15 word maximum, and cons with 15 word maximum, it's something annoying because maybe the user doesn't want to write any pros or cons.

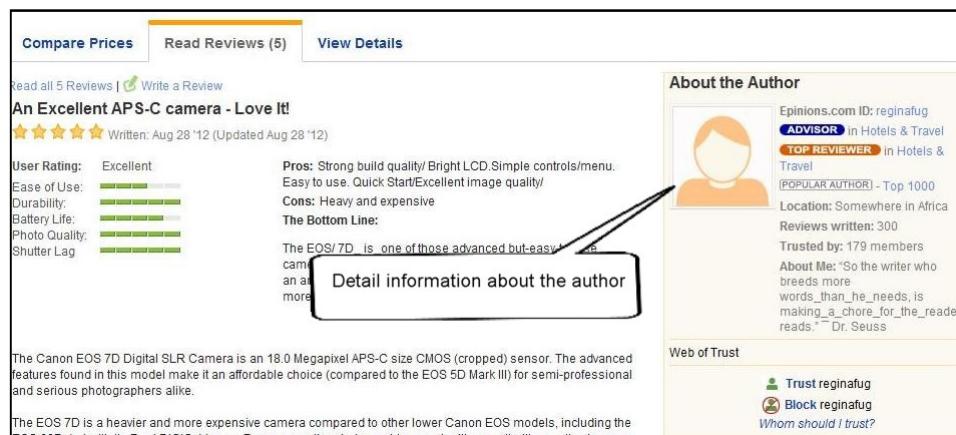


Figure 50: Epinions provides detail information about the author

The Epinions system helps users to find out what additional steps are needed in order to complete a task successfully e.g. if the user writes a review with fewer than 200 words, the system will provide a message explaining the situation as shown in figure 51, and if errors do occur, the system will provide a user friendly message in plain language, e.g. if the user wants to leave a review and he forgets one of the fields, a user friendly message will appear; "The following is required." As shown in figure 52.

Epinions has a moderate help system, in some places it provides an accurate help, and sometimes you can't find a help so you need to search about it.

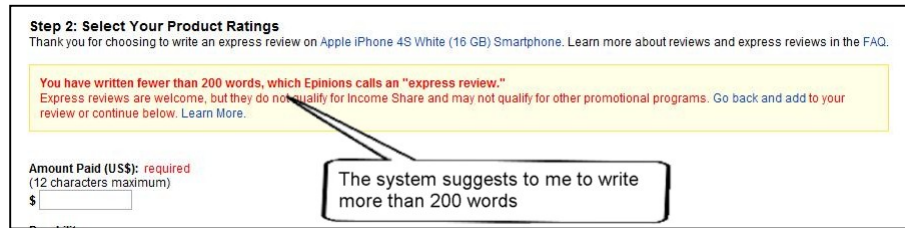


Figure 51: Epinions provides suggestion to users

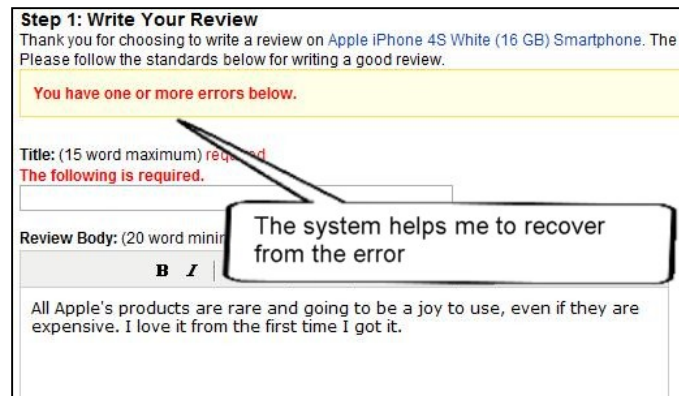


Figure 52: Epinions provides friendly message when error occurs

#### 4.1.6 iTunes Apps Store

The App Store is an online Apple Store that allows users to download and/or buy application that runs on Apple iPhone/iPad running the iOS operating system. The App Store functions much like the iTunes Store, and is integrated with iTunes to deliver content to devices and for purchases. The same iTunes account used for iTunes purchases is used for the App Store.

The applications on the App Store are generally written by third-party developers (Apple offers a few of its own apps) and submitted to Apple for approval. Apple uses an extensive review process to check the apps for adherence to the company's guidelines and programming/content standards. If Apple approves the program for inclusion on the store, it is published to the site, where users can download/purchase it.

Developers rate their own applications either with 4+, 9+, 12+, or 17+ levels, which refers to age restrictions. The ratings are as follows (from iTunes):

4+: Applications in this category contain no objectionable materials.

9+: Applications in this category may contain mild or infrequent occurrences of cartoon, fantasy or realistic violence, and infrequent or mild mature, suggestive, or horror-themed content which may not be suitable for children under the age of 9.

12+: Applications in this category may also contain infrequent mild language, frequent or intense cartoon, fantasy or realistic violence, and mild or infrequent mature or suggestive themes, and simulated gambling which may not be suitable for children under the age of 12.

17+: You must be at least 17 years old to purchase this application. Applications in this category may also contain frequent and intense offensive language; frequent and intense cartoon, fantasy or realistic violence; and frequent and intense mature, horror, and suggestive themes; plus sexual content, nudity, alcohol, tobacco, and drugs which may not be suitable for children under the age of 17.

Customer reviews are one of the main tools in the apple applications, where people who've downloaded a certain app write a little paragraph about it and give it a star rating, 1 through 5. At a glance, any user can find out whether a certain program is a good or not, but the problem with the rating is sometimes a software company deliberately trashes the apps of its competitors, giving them terrible reviews in hopes of damping their sales.

The most important points in the in the iTunes apps Store's feedback profile figure 53 are:

1: (General Application Information) details information about the app, e.g. version, size and developer name.

2: (Specific Application Information) details information about the app, e.g. what are the features of this app.

3: (Rated) it shows the app rated number, which refers to age restrictions.

4: (Requirements) it shows the system requirement that are compatible with the app.

5: (Developer other apps) it shows other apps by the same developer.

6: (Current Version Star Rating): the average rating for the current app version in all the time.

7: (All Versions Star Rating): the average rating for the app all versions in all the time.

8: (Current Version Customer Reviews): this refers to the reviews that are left from the users to the current version.

9: (All Version Customer Reviews): this refers to the reviews that are left from the users to the all versions. [11]

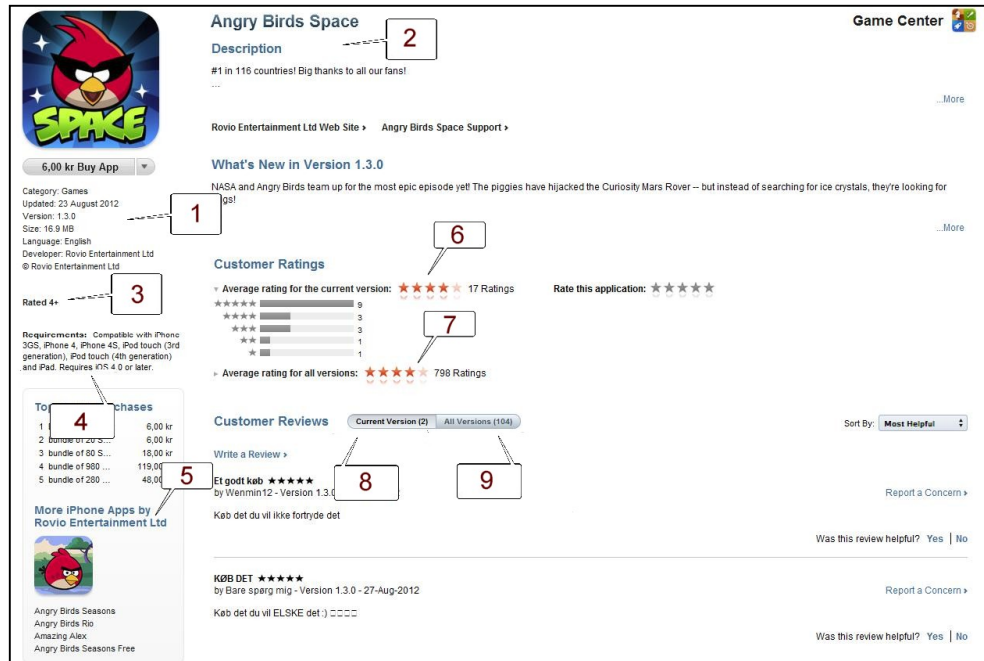


Figure 53: iTunes app Store rating system

## Nielsen heuristics for iTunes App Store' Reputation Systems

There are some negative and positive points in the iTunes App Store reputation system depending on our Nielsen heuristics evaluation. The most important points are; the user always knows where he is and what he is doing in terms of carrying out a procedure using a system figure 54, Users would know if an operation was successfully completed (e.g., Your review has been successfully submitted.) figure 55.

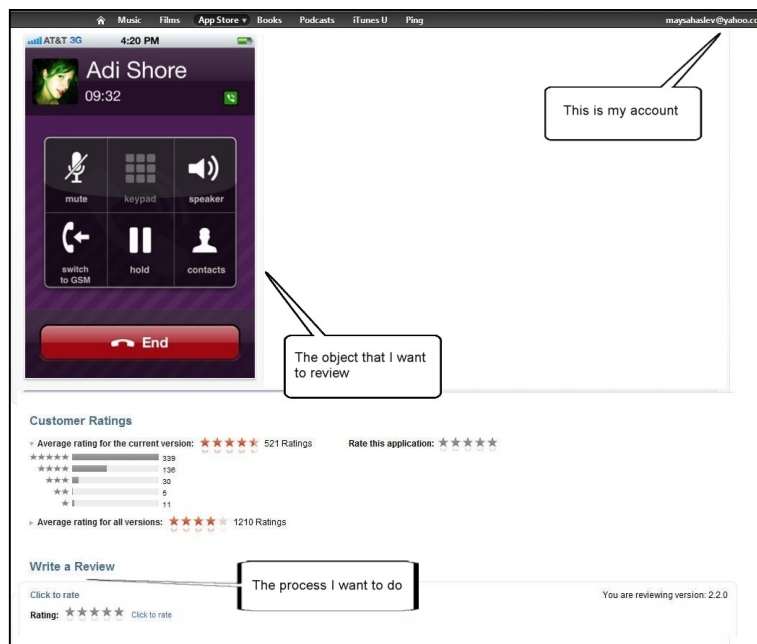


Figure 54: iTunes visibility

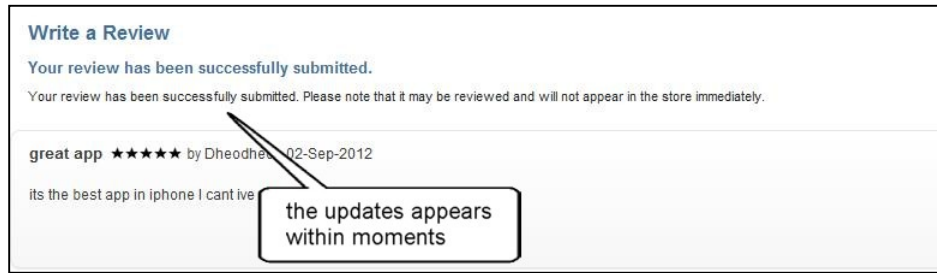


Figure 55: iTunes updates within moments after review submission

iTunes App Store reputation system uses natural words; particularly novice users can understand the language. One of the positives in the iTunes App Store reputation system is; the system provides the option to edit the review and the rating or delete the review he left earlier this will allow the user the possibility to correct any mistakes that have been made before or maybe the user changes his mind about the reviewed app, this will add more trustworthiness to the ratings as shown in 56.

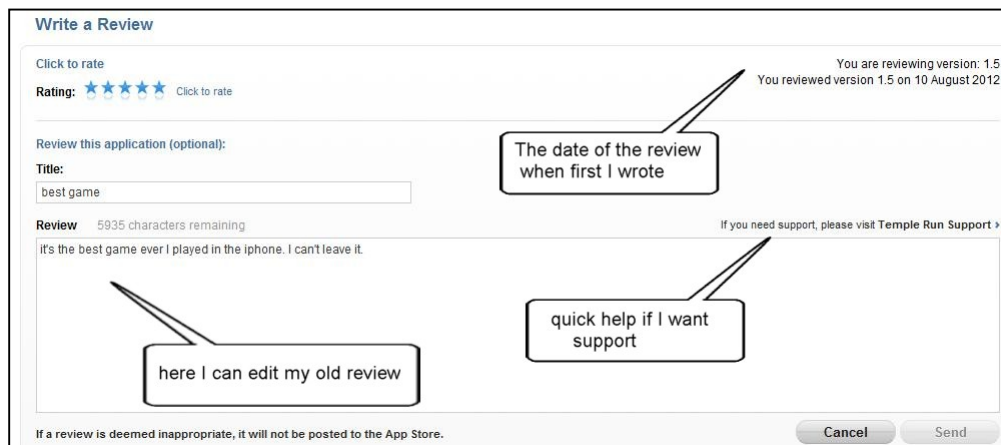


Figure 56: iTunes Control and freedom

The Star rating system in iTunes App Store has some negatives such as the user has no information about the way the ratings are calculated, it just appears that they are based on the reviews, but on the other hand there are some positives such as the iTunes App Store system explains clearly what each star means before we rate the app as shown in figure 57.



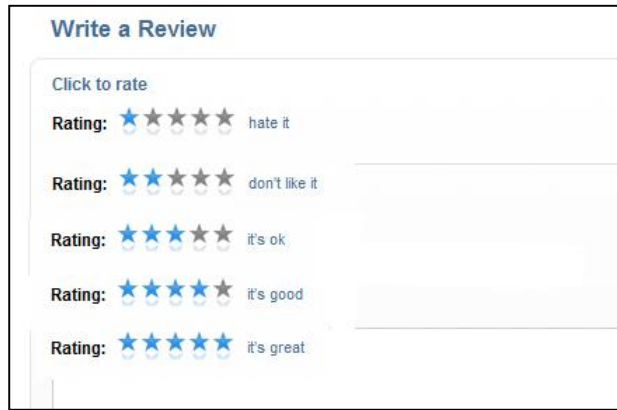


Figure 57: clear understanding to the star rating in iTunes

In the iTunes App Store system the users can retrieve all the history of reviews for the app current version and the app all other versions, so one can compare the current version with previous versions. There is also a possibility to rate a review with a helpful or to report it. This will add trustworthiness to the reviews figure 58. One of the negatives in the iTunes App Store system there isn't enough information about the reviewers, so we can't know if this reviewer is trustworthy or not.

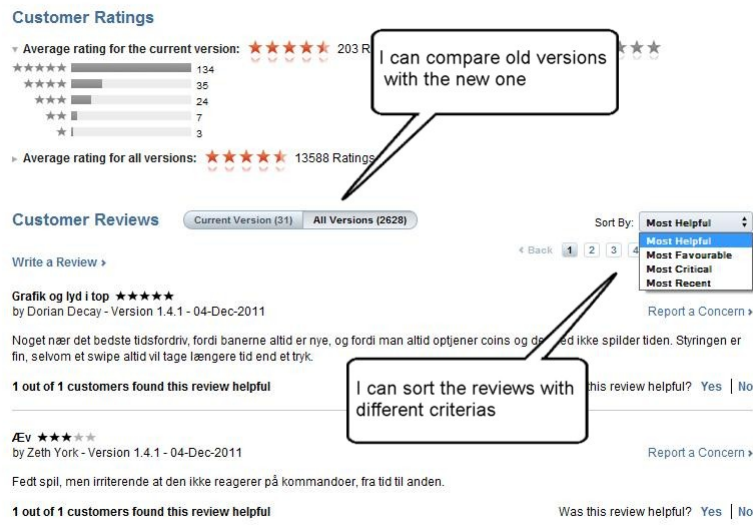


Figure 58: iTunes provides sorting reviews by different criteria with comparison

The iTunes App Store system provides some facilities that help users to speed up their interaction with the system such as "View More By This Developer"; to visit other apps by the same developer as shown in figure 59.



Figure 59: iTunes provides some facilities

iTunes store system provides a detailed overview about each app, so the user can know everything about this app before he installs it. When a user clicks on an app, all information will be provided in two sections – 1) a description with snapshots, and 2) ratings and reviews by users. It provides the possibility to sort the review depending on different criteria: most helpful, most favorable, most critical and most recent as shown in figure 58.

The iTunes App Store system helps users to find out what additional steps are needed in order to complete a task successfully, and if errors do occur, the system provides a red message that helps the user to recover from the error, e.g. if the user wants to leave a review and he forgets to rate the app, a message in red will appear; “Click to rate” as shown in figure 60.



Figure 60: iTunes helps users to recover from errors

iTunes App Store doesn't provide a good help system, if you have a question or don't understand something you must search about it without any help from the system, but sometimes it offers you some support as shown in figure 56.

#### 4.1.7 Youtube.com

You tube is a video sharing website, where users can upload, view and share videos. Unregistered users can watch videos, while registered users can upload an unlimited number of videos. Videos considered to contain potentially offensive content are available only to registered users at least 18 years old. All YouTube users can upload videos up to 15 minutes each in duration. Users who have a good track record of complying with the site's Community Guidelines may be offered the ability to upload videos up to 12 hours in length, which requires verifying the account, normally through a mobile phone. YouTube has changed its Reputation System from a 5 star to Thumbs-up/Thumbs-down ratings. In the most recent design, the video being played shows the number of likes and dislikes. Most videos enable users to leave comments and enable other users to rate the comments by voting up or down. Only the registered users can rate a video and write comments. [12]

The most important points in the in the YouTube's reputation profile figure 61 are:

- 1: (Like Ratings) the number of likes and dislikes depending on the viewer action.
- 2: (Statistical Information) details statistical information about the video, e.g. number of comments, top locations the video watched from and top demographics that watched the video
- 3: (Video Uploader Information) one can see the uploader name and the date he uploaded the video.
- 4: (Video General Information) details information about the video e.g. information about the uploader and information about the event of the video.
- 5: (Top Comments) the comments that has high numbers of thumbs-up.
- 6: (All Comments) all the comments which is sorted by date.

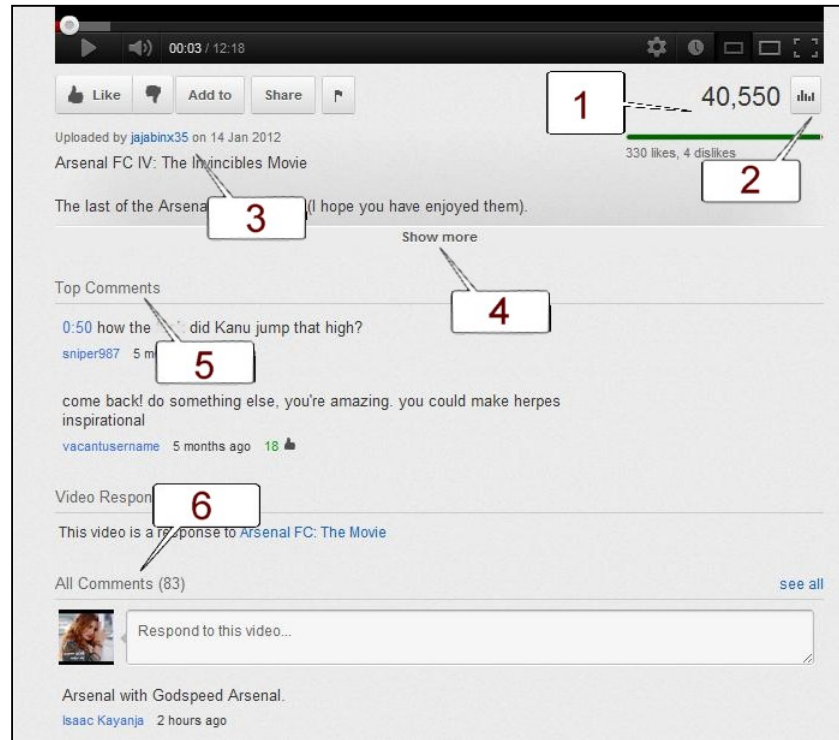


Figure 61: YouTube Rating Profile

## Nielsen heuristics for YouTube's Reputation Systems

There are some negative and positive points in the YouTube reputation system depending on our Nielsen heuristics evaluation. The most important points are; the user always knows where he is and what he is doing in terms of carrying out a procedure using a system figure 62, users would know if an operation was successfully completed e.g. when a user clicks on like icon under the video; a message will appear "Thanks! Share it with your friends" figure 63.

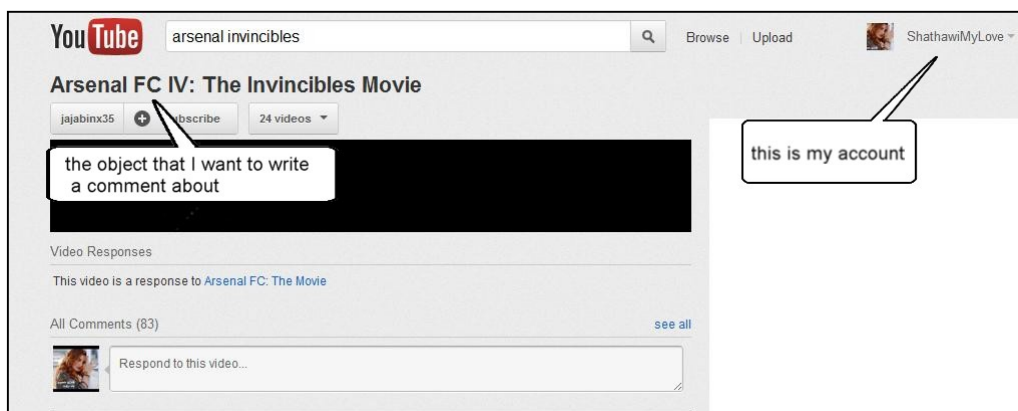


Figure 62: YouTube showing account name and the reviewed video



**Figure 63: YouTube provides the updates within moments**

YouTube reputation system uses natural words; particularly novice users can understand the language and the system categories are well grouped. One of the negatives in the YouTube reputation system is; there is no redo/undo mechanism in some fields. The user can undo the action of the rating a video, but cannot undo the action of reporting another user. This will add unreliability in the reputation information, because users can report and be reported mistakenly. The user has the possibility to remove his comments on a video as shown in figure 64.



**Figure 64: YouTube provides only remove choice to my comments**

There is no absolute rating for the video; the rating is Thumbs-up/Thumbs-down ratings, but the showing the number of likes and dislikes for each video would improve the quality of videos on YouTube by discouraging people from posting bogus or hoax videos, and also by promoting good videos and allowing more viewers to find quality content on YouTube. The recommended Videos column on the right displays only thumbnails and the number of views. So a user has no way to know the authenticity of the video without actually clicking on it, figure 65.

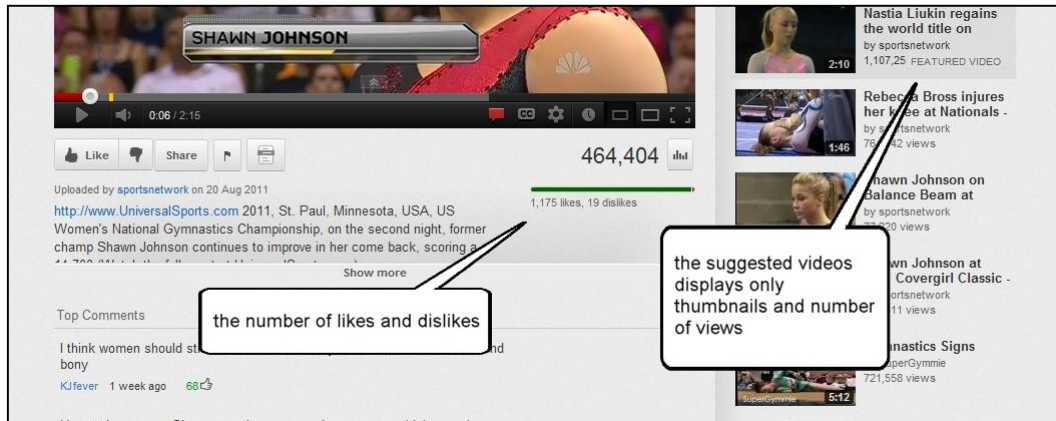


Figure 65: YouTube like/dislike system

In the YouTube system the users can retrieve all the comments of other users and the top comments that have been rated highest, but the replies to the video comments are not displayed close to the message that the reply was connected to; just the reply appears at the top, and if you want to see the original comment you must press on the "show the comment".

The YouTube system provides some facilities that help users to speed up their interaction with the system such as "Show the comment" as shown in figure 66, and It's easy to rate a video while you are watching it.

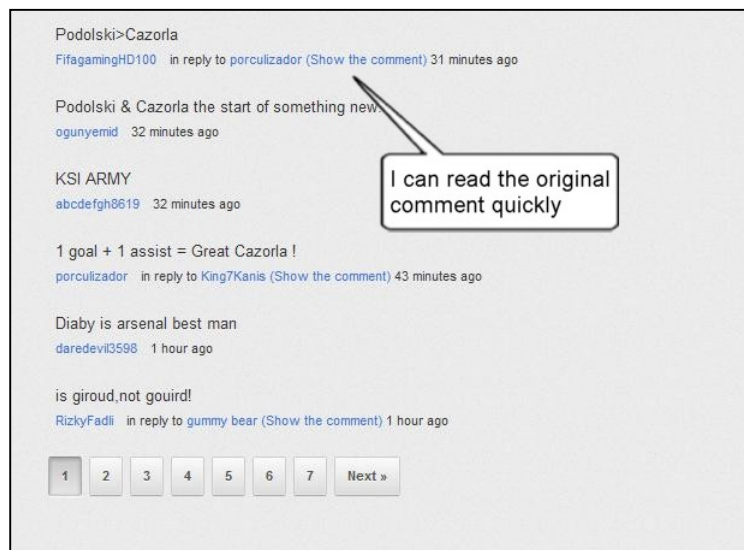


Figure 66: YouTube provides some facilities that speed up the interaction with the system

There are other features that add some efficiency in the YouTube, for example the user can hover his mouse over the seek bar and a thumbnail of that moment will appear, and he can click to start watching from that moment figure 67.



Figure 67: YouTube taking a sneak peek

Another feature in the YouTube is, when for example you left a video in the middle and you want to continue from that point, you can drag the handle along the seek bar to show a filmstrip of thumbnails of pervious and upcoming scenes figure 68.

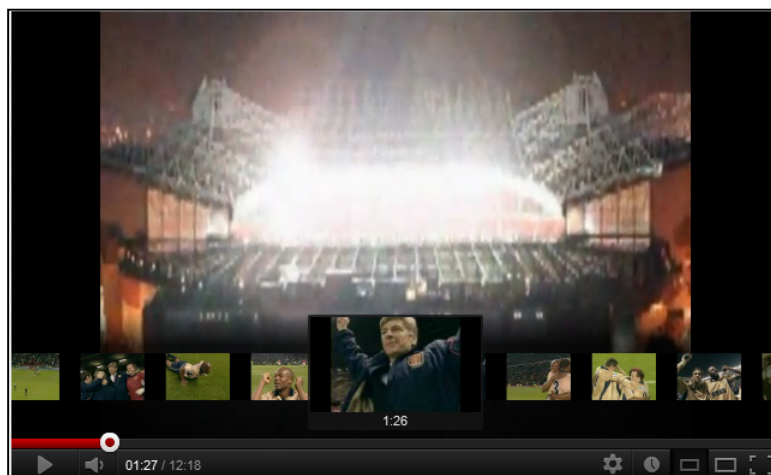


Figure 68: YouTube scanning through thumbnails

Another nice feature in YouTube if a user watch a video that is longer than 90 minutes, he can see an added feature that lets him zoom in on the seek bar, one and a half minutes at a time. The second bar that appears gives the user the aid for finding that exact moment he wants on a long video figure 69. The user can also choose the quality of the video as shown in figure 69.



Figure 69: YouTube zooming in one long video

There is a plenty of information about the video under the video and about the user who uploaded the video. There is a very good specific statistical information about each video, which shows the history, popularity and the number of viewers of the videos figure 70, but the comments are sorted only by date in spite of they choose some top comments, which they are published in front of the other comments.

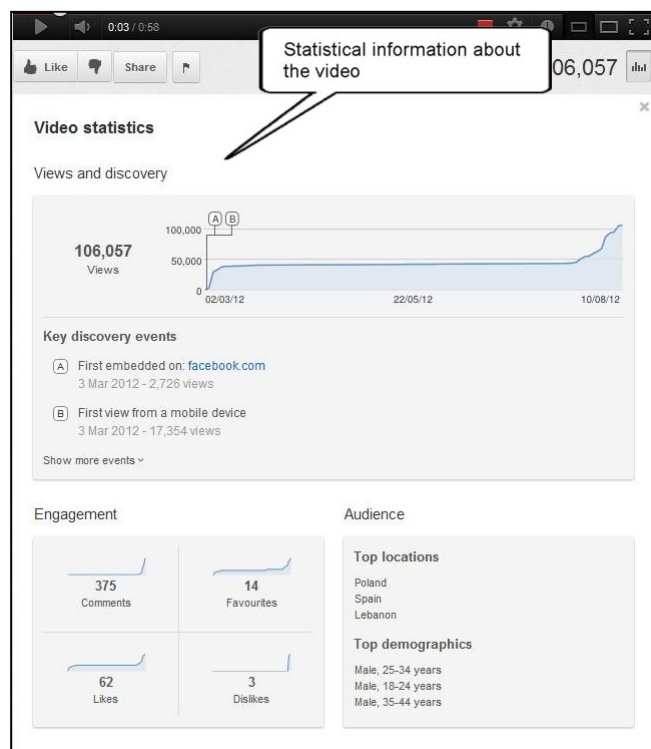


Figure 70: YouTube provides statistical information about any video

The YouTube system will show a message when something wrong happens as shown in figure 71, but this sometimes will lead to confuse the users, because the message is not



friendly; there is no explanation what is the cause of the error. YouTube provides help to user when they need by pressing on help at the end of the page.



**Figure 71: YouTube provides help to recover from errors**

## 4.2 Applications of Think aloud and Remote Testing

In this section we will present the results of the think aloud and remote testing that is conducted by asking six people to test the chosen sites.

### 4.2.1 eBay.com



All the test participants said that they understand clearly how the reputation system of the eBay works, and how the seller percentage is calculated.



All the test participants said that they would contact the seller before leaving a natural or bad review. But if the review was good they will leave a feedback directly.



All the test participants commented favorably on the website graph. They thought that the graph was suitable and didn't disturb them and gave them a good idea of what are the important areas in the website.



All the participants said that the navigation was clear and it was very easy to navigate the website.



All the participants appreciated that they can find all the relevant information about the seller and the items, and they can find the information they needed easily.

### 4.2.2 Amazon.com



Two of the six test participants said that they understand how the reputation system of Amazon works, and how the seller percentage is calculated, but it would be a good idea if there is more explanation about the seller reputation.



Four of participants annoyed that they could not understand how the reputation system works, and how the seller percentage is calculated, they couldn't differentiate the stars what refer to, whether to the quality of the items or shipping service or low price.



All the test participants said that they would contact the seller before leaving a natural or bad review. But if the review was good they will leave a feedback directly.



Three of the test participants commented on the website graph by saying it's ok. They thought that the graph was suitable.



Three of the six test participants suggested redesigning the website because they thought that the design and graph of the website didn't suit a big and famous e-commerce, so the site needs a new design.



Five of the test participants found the navigation of the Amazon not easy and there are so many choices, which get them dizzy. It was difficult to them to find what they search for.



Two of the test participants said they could find information about the seller and the items, but it was difficult to them to find the negative reviews about the seller.



Three of participants were confused and find it difficult to find the negative and natural reviews about a seller. They said:

*"Oh I'm very confused, I don't know where to go, I just give up".*



Four of the test participants thought the retrieving of the seller information is difficult, and couldn't find the information about the seller easily.

#### 4.2.3 **Yelp.com**



Five of participants find it easy to understand clearly how the reputation system of the Yelp works.



None of the participants could understand how the rating score is calculated.



Four of test participants said they would leave review whether good or negative review and it's easy to leave a review.



Two of participants found it difficult to leave a review



All the test participants commented favorably on the website graph. They thought that the graph was suitable and didn't disturb them and gave them a good idea of what are the important areas in the website.



All the participants said that the navigation was clear and it was easy to navigate the website.



Five of the participants appreciated that they can find all the relevant information about the restaurant, and they can find the information they needed easily.



One of participants found it difficult to find the negative reviews about a restaurant.

#### 4.2.4 **TrustPilot.com**



Five of the test participants said that they understand clearly how the reputation system of the trustpilot works.



None of the participants could understand how the rating score is calculated.



All the test participants said that they can easily write a review about a company.



All the test participants commented favorably on the website graph. They thought that the graph was suitable and didn't disturb them and gave them a good idea of what are the important areas in the website.



All the participants said that the navigation was clear and it was ok to navigate the website.



None of the participants could categorize the negative reviews about a company.

#### 4.2.5 **Epinions.com**



Three of participants find it easy to understand clearly how the reputation system of the Epinions works.



Three of the test participants couldn't understand clearly the reputation system and was not sure what the rating means.



None of the participants could understand how the rating score is calculated.



Two of participants find the review system is ok and helpful to the customers.



Four of the six test participants didn't like the review system and suggested to redesign the review system because they thought that the review system requires too many fields and that is boring.



Five of participants liked the website graph. They thought that the graph was suitable and didn't disturb them and gave them a good idea of what are the important areas in the website.



One of the participants didn't like the graph of the website and found it hurtful to eyes.



All the test participants found the navigation of the Epinions difficult to navigate and they get dizzy. It was difficult to them to find what they search for.



Four of the test participants found it difficult to retrieve information about a product.



Two of the participants appreciated that they can find all the relevant information about the product, and they can find the information they needed easily.

#### 4.2.6 **ITunes Apps Store**



Two of the test participants said that they understand clearly how the reputation system of the ITunes app store works.



Four of participants annoyed that they could not understand how the reputation system works.



None of the participants could understand how the rating score is calculated.



Two of participants they didn't understand the rated criteria; they search for help to understand it.



All the test participants said that they can easily write a review to any application.



All the test participants commented favorably on the website graph. They thought that the graph was suitable and didn't disturb them and gave them a good idea of what are the important areas in the website.



All the test participants said that the navigation was clear and it was ok to navigate the website.



Two of the participants appreciated that they can find all the relevant information about the developer and the applications, and they can find the information they needed easily.



All the participants found it easy to find the critical reviews about an application.

#### 4.2.7 **YouTube.com**



All the test participants said that they understand clearly how the reputation system of the YouTube works, and how the seller percentage is calculated.



All the test participants said that they can easily write a review to any video.



Three of the participants annoyed that they couldn't remove their previous like to a video.



Four of the test participants commented favorably on the website graph. They thought that the graph was suitable and didn't disturb them and gave them a good idea of what are the important areas in the website.



Two of the six test participants didn't like the graph and suggested to redesign the graph of the website, they thought that it's too mess and that is confusing.



All the participants said that the navigation was clear and it was very easy to navigate the website.



All the participants appreciated that they can find all the relevant information about the video and the video's uploader, and they can find the information they needed easily.

## 4.3 Usability of the reputation system of the chosen Websites

It has been referred in section 2.4 that the usability is associated with five usability attributes: learnability, efficiency, memorability, error handling and satisfaction. So according to our evaluation using the Nielsen ten heuristics and the think aloud testing findings, the five usability attributes for the seven reputation systems will be discussed in the following sections.

### 4.3.1 eBay

**Learnability:** the novice users of the eBay don't need time and training before they become capable of understanding and using the eBay's reputation system. The system offers help and many facilities which guide the users to do the required process. And when a user needs a help he can use the help which is available in the same place. The help document contains detailed information about the ambiguous icon or link. All the categories and menus in eBay are well grouped. So we can consider eBay doesn't have a problem with learnability.

**Efficiency:** all the experienced users were able to complete the given tasks quickly without any problem. So the eBay has an efficient working process.

**Memorability:** the users who use the eBay infrequently, they can remember the use of the system easily. The help offered with each step and details information about each icon and link in the eBay has make it easy to recall for an infrequent user. So the system manages quite well when it comes to helping the user in recollection of its use, mainly due to a clear working process and the detailed help document.

**Error:** One of the negatives in the eBay reputation system is; the system does not contain the undo functions, so if the user leaves a feedback and he regrets it, he cannot edit it or remove it, so this will make the system difficult to recover from previous rating and feedback. But the eBay system provides a good feedback to users to find out what additional steps are needed in order to complete a task successfully, or what they are did wrongly to resolve the errors on their own. So based upon the previous evaluation on eBay, we think the eBay system's error handling should be better.

**Satisfaction:** all the test participants were very pleasant and satisfied with the eBay in spite of the lack of the undo function in the leaving feedback process. They all expressed their desire to use eBay again. So we think that the user satisfaction in eBay is high.



### 4.3.2 Amazon

**Learnability:** the novice users of the Amazon will need time and training before they become capable of understanding and using the Amazon's reputation system; this is because of the similarity in representation of the reputation information. The novice users get confused with the product description and other information. They don't know where to find the information they search about it e.g. about dispatch countries, but on the other hand there are some facilities and helps the Amazon offers; the help document contains detailed information about different subjects, the users can press on these links to find a quick help about them. All the categories and menus in Amazon are well grouped. So it is expected that the novice user should be given some training before they can do a given task, but it has many good features that help the novice users to quickly learn to use and understand the system. So Amazon has some problems that need to be fixed.

**Efficiency:** some of the experienced users have encountered some problems in a given task, for example they take a time to figure out that the item they want to buy can't be shipped to the selected destination. So the system needs to be improved by informing the buyers the countries that the seller will ship to. Another problem in Amazon most of the users need to read the negative reviews about a seller, before they want to purchase from him, but this feature is not available in Amazon. The user needs to navigate many pages until he may reach to a negative review. So based on the above problems, it can be stated that some of Amazon's problems needs to be fixed to be more efficient.

**Memorability:** the infrequent users can recall the Amazon easily. The help in Amazon are offered in many places with details information about each icon and link. The detailed help document has helped the infrequent users to remember how to use the system easily. So we can consider Amazon has no problem with memorability.

**Errors:** One of the negatives in the Amazon reputation system is; the system does not contain the undo mechanism, so the user cannot change or comment his earlier feedback, although he can remove the negative ones but once he removes it he can't post a new feedback instead of the removed one, so this will make the system difficult to recover from previous rating and feedback. But the Amazon system provides a user friendly feedback to users to find out what additional steps are needed in order to complete a task successfully, or what they are did wrongly to resolve the errors on their own. So based upon the previous evaluation on Amazon, we think the Amazon system's error handling should be better.

**Satisfaction:** half of the test participants were pleasant and satisfied with the Amazon reputation system, but the other half were irritating and unsatisfied with it. They found it confusing to take a decision or navigate about the user's profile. The lack of the possibility

to filter the negative reviews makes it inconvenient to use. So half of the participants expressed their desire to use it again, and the other half ruled out to use it again. So we think that the user satisfaction in Amazon is middle.

### 4.3.3 Yelp

**Learnability:** the novice users of the Yelp don't need time and training before they become capable of understanding and using the Yelp's reputation system. The yelp system provides help and facilities which guide the users to do the required process. The help document in yelp is ok, when a user needs a help he can use the help that is available in the same place, but sometimes the user cannot find a quick help so he needs to search about it. We can consider Yelp doesn't have problem with learnability.

**Efficiency:** all the experienced users were able to complete the given tasks quickly without any problem. So the Yelp has an efficient working process.

**Memorability:** the nice overview of yelp makes it easy for infrequent user to remember the process that they want to use. The Yelp's interface is built on the principle of making as much as possible visible to users. The system has no problem with memorability.

**Errors:** One of the positives in the yelp reputation system is; the system provides the "undo" functions. The users can edit or remove the ratings and the reviews they did before to the business; this will allow the user to recover from any action he did earlier. The yelp system helps users to find out what additional steps are needed in order to complete a task successfully, and if an error does occur, the system provides a feedback to help the users to resolve the error on their own. So based on these facts, we can find that the Yelp system has a good error handling.

**Satisfaction:** most of the test participants were pleasant and satisfied with the Yelp reputation system, but just one of the participants was irritating and unsatisfied with it. He found it difficult to filter the negative reviews and difficult to leave a review. He found it complicated to use. So most of the participants expressed their desire to use it again, and two of them ruled out to use it again. So we think that the user satisfaction in Yelp is ok.

### 4.3.4 Trustpilot

**Learnability:** the novice users of the Trustpilot don't need time and training before they become capable of understanding and using the trustpilot's reputation system. The

system is very easy and the new user can use the system without any difficulties. The help document in trust pilot is moderate; in some places the user can find an accurate help for example by pressing on an icon, and sometimes the user can't find help so he needs to search about it. So we can consider Trustpilot does not have a problem in learnability.

**Efficiency:** all the experienced users were able to complete the given tasks quickly without any problem. So the Trustpilot has an efficient working process.

**Memorability:** the simple overview of Trustpilot makes it easy for infrequent user to remember the process that he wants to use. The system interface is very simple and the process of using Trustpilot is very clear, so the infrequent user can recollect the use of the system when he wants to use it again. The system fares quite well when it comes to helping the user in remembrance of its use, mainly due to the clear working process.

**Errors:** One of the positives in the Trustpilot reputation system is; the system provides the option to edit or delete the reviews he left earlier, this will allow the users to correct any mistakes have been made before or maybe they change their mind about the reviewed company. The Trustpilot system helps users to find out what additional steps are needed in order to complete a task successfully, and if an error does occur, the system provides a user friendly message in plain language to help the user to resolve the error on his own. So based on these facts, we can find that the Trustpilot system has a good error handling.

**Satisfaction:** four of the test participants were pleasant and satisfied with the Trustpilot reputation system, but the other participants were irritating and unsatisfied with it. The lack of the possibility to sort the reviews by "rating" makes it inconvenient to use. So half of the participants expressed their desire to use it again, and the other half ruled out to use it again. So we think that the user satisfaction in Trustpilot is middle.

#### 4.3.5 **Epinions**

**Learnability:** the novice users of the Epinions will need time and training before they become capable of understanding and using the Epinions's reputation system; this is because the navigation is not easy; furthermore when they want to leave a review, they think it is boring and it requires so many fields. Another difficulty is, they don't understand the difference between the author's status, e.g. advisor and top reviewer, they cannot find a quick help, so they need to search about what the difference is between the author's statuses. So the help system is moderate. In some places it provides accurate help and in others can't provide it. We can consider Epinions has some weaknesses and needs to be fixed.

**Efficiency:** some of the experienced users have encountered some problems in a given task, for example when they want to write a review they take time to read the instruction on the write of the page to figure out what they should avoid in writing a review. Another problem in Epinions that faces some experienced users when they sort the product by rating, it appears that the item has 1 review and when the expert want to read this review he cannot find it, that's may be due to the delay of the updating of the system. So based on the above problems, it can be stated that Epinions needs some fixing to its problem to be more efficient.

**Memorability:** the infrequent users who are using the Epinions intermittently, they have not a problem to recollect the use of the system. In spite of the overview of the system is not simple, the users who have used it before they can remember the use of this system. There is no problem in the system in regarding the memorability.

**Errors:** One of the positives in the Epinions reputation system is; the system provides the option to undo (update, move or delete) the ratings and the reviews the user did before to the product; this allow the user to modify his review if he has changed his mind or correct any mistakes that have been made. The Epinions system helps users to find out what additional steps are needed in order to complete a task successfully and if an error does occur, the system provides a user friendly message in plain language to help the users to resolve the error on their own. So based on these facts, we can find that the Epinions system has a good error handling.

**Satisfaction:** most of the test participants were irritating and unsatisfied with the Epinions reputation system. They describe the process of leaving a review is a boring task and a repetitive task of inputting so many words in just one review. The difficult navigation and the confusing about the products make the system inconvenient to use. Most of the participants rule out to use the system again. So we think that the user satisfaction in Epinions is low.

#### 4.3.6 iTunes Apps Store

**Learnability:** the iTunes system is not easy and the novice users will need time and training before they become capable of understanding and using it; this is because they need to know the iTunes application, which is not an easy application. So considering carrying out a given task to a new user, it is expected he needs some training. The help system is not good in iTunes system; you should search about any problem you encounter. There is no quick help in the system. So we can consider iTunes has some issues in learnability that need to be considered.

**Efficiency:** all the experienced users were able to complete the given tasks quickly without any problem. So the iTunes apps store has an efficient working process.

**Memorability:** the nice overview of the iTunes has made it easy for an infrequent user to recognize and use the system easily. The system doesn't have a problem with its memorability.

**Error handling:** the iTunes system provides the option to edit and delete the review, while the rating can only be changed but not removed. The iTunes App Store system helps users to find out what additional steps are needed in order to complete a task successfully, and if an error does occur, the system provides a red message (not user friendly) that helps the user to recover from the error. So based on these facts we think that the iTunes error handling could have been better.

**Satisfaction:** all the test participants were very pleasant and satisfied with the iTunes app store. They all expressed their desire to use it again. So we think that the user satisfaction in iTunes app store is high.

#### 4.3.7 YouTube

**Learnability:** the novice users of the YouTube don't need time and training before they become capable of understanding and using the YouTube reputation system. YouTube is very easy and offers help and facilities which guide the users to do the required process. And when a user needs a help he can use the help that is available in the same page. There is no problem with learnability in YouTube.

**Efficiency:** one of the experienced users had a problem when he wanted to unlike an old liked video; he spent a good time until he could figure out how to do this action. So we can consider the YouTube needs some improvement to be more efficient.

**Memorability:** the simple and clear overview of the YouTube has made it very easy to remember for the infrequent users. So the system doesn't have problem with memorability.

**Error handling:** One of the negatives in the YouTube reputation system is; the user cannot undo the action of reporting another user, and it's very difficult to remove a video that previously rated as "Liked" from the "liked videos", but the user has the possibility to remove his comments on a video. The YouTube system shows a message when an error does occur e.g. "Error. Try again" and this sometimes will lead to confuse the users, because the message is not friendly; there is no clear message what the cause of the error is. So based on these facts we think that the YouTube error handling could have been better.

**Satisfaction:** all the test participants are very pleasant and satisfied with the YouTube in spite of some of them don't like the new overview of the YouTube. They all expressed their desire to use it again. So we think that the user satisfaction in YouTube is high.

## 4.4 Applications of Survey and Interviews

After we finished the previous usability inspection (Heuristic evaluation) and usability testing (think aloud testing) and based upon the usability summary in the previous section, where each of the five usability attribute measures has been discussed, the conclusion was to choose the best, natural and worst usability among the seven websites.

These sites are:

- eBay (the best usability)
- Amazon (natural usability)
- Epinions (the worst usability)

After the previous work, we began another usability study (usability inquiry) by interviewing website based-reputation systems' users to gain an initial understanding on how reputation systems are currently perceived and used. The interviews and the survey are based on the three filtered website that are chosen from the previous evaluation.

A web-based survey\* was built on basis of the interview outcomes to gather feedback from more users. The questions in the survey consisted of a modified version of the interview questions. We had in all more than 100 users but we filtered and took the result of only 56 users who have used the three sites at least one time.

78 % had at least a bachelor's degree. All interviewees (n=8) were different nationalities located in Denmark, whereas the survey respondents represented various other nationalities.

The results of the interviews and the survey confirmed the following:

1. Over 90% of the users understand clearly the rating systems for the three sites and 40% of them think that eBay and Amazon have an over average understandable rating system, while just 25% of the users think the Epinions has an over average rating system.
2. 90% of the users think that eBay provides enough information about the seller, while 70% for Amazon and Epinions.
3. 80% of the users think the three sites provide all the needed information at one glance.

<http://obsurvey.com/S2.aspx?id=E624A728-FF9B-4C11-ABEC-4236E4A322E8>

4. 90% of the users think that eBay provides all the information they need to decide, while 70% for Amazon and Epinions.
5. 60% of the users find eBay easy when they use it and just 1% of the users find it difficult to use, while 50% find Amazon easy to use and 10% find it difficult to use. For Epinions there are 30% of the users who find it easy and almost half of the users find it difficult, these were the answers of the users when they were asked about how easy the three sites to use.
6. When the users asked whether they are satisfied with the three sites services the answers were: 80% of users express their satisfaction with eBay services with very satisfied and satisfied and no answer with unsatisfied. While for Amazon there were 60% of users express their satisfaction with the site services and 25% express their dissatisfaction. For Epinions there were 27% of users are satisfied with Epinions services and almost half of the users are unsatisfied.
7. When the users asked what do you like and don't like about eBay: According to the results from this set of participants, the three entities that they rate the most are: 50% of them think the site is very easy to use, 40% think there are detailed information about the item, and 35% of them think the language of the site is understandable. While the three entities that they don't like about eBay even if 50% they said there is nothing in eBay they don't like are: 15% of users think there is so many information and features that may confuse them, 12% for the both there is no undo in the feedback system and the user can only sort reviews by date.
8. When the users asked what do you like and don't like about Amazon: According to the results from this set of participants, the three entities that they rate the most are: 45% of them think that there is detailed information about the items, 33% for both the language of the site is understandable and the site is easy to use. While the three entities that they don't like about Amazon even if 45% they said there is nothing in Amazon they don't like are: 20% of users think there is so many information and features that may confuse them, 13% of users think there is no enough information about the seller, and 11% of users think there is no more information about the items; e.g. the users would like to see the inside of a book before they purchase it.
9. When the users asked what do you like and don't like about Epinions. According to the results from this set of participants, the three entities that they rate the most are: 22% think the site is easy to use, and 20% for both the site provide detailed information about the items and detailed information about the compared prices.



While the three entities that they don't like about Epinions are: 22% of users think it's difficult to navigate the site, 18% of users think the site is hard to use, and 14% of users think there is no more information about the items.

10. When the users asked whether they have suggested the sites to others, 70% of users have suggested eBay and Amazon to others, while only 17% have suggested Epinions to others.
11. 85% of the users read the reviews before they buy a product and 50% of users contact neither the seller nor other buyers from the seller before they buy a product.
12. When the users were asked about how often they write a feedback 40% of users said that they would never leave negative or natural feedback, while 40% will leave feedback if it was positive.
13. To find out what factors (reputation elements) were taken into account when they want to buy a product, we presented to the users screenshots of the different reputation elements; the answers are: 72% of users think the product picture and description is the most important factor, the second factor was the percentage of the number of positive feedback with 65% and 58% think the seller star rating is the third important factor. The latest feedback review is the fourth factor with 42% when they want to buy a product.
14. To find out what factors (reputation elements) were most of the users don't take into account when they want to buy a product, we presented to the users screenshots of the different reputation elements; the answers are: 40% of users think the communication rating is not important, 37% of users think postage and packaging charges rating is the second not important factor, while 28% of users think the dispatch time rating is the third not important factor.
15. To find out what factors (reputation elements) were the users don't understand them, we presented to the users screenshots of the different reputation elements; the answers are: 35% of users don't understand what the communication rating means, 20% of users they don't understand the top-rated seller icon and dispatch time rating.
16. When presented snapshot with examples of same product from eBay, Amazon and Epinions to choose a site to buy the product if the reputation of the seller is the most important factor to you, 70% of the participants consider eBay, 24% will choose Amazon and 6% consider Epinions.

17. When presented with examples of eBay, Amazon and Epinions feedback mechanism to write a review, 95% of the participants consider writing a review in eBay and Amazon, while only 5% consider Epinions is very boring and difficult to write a feedback.

# Chapter 5

## 5. Discussion and Recommendations

---

In this chapter the data will be discussed with the research questions. There are some recommendations that will be proposed to help improving the usability of the reputation-based websites.

### **5.1 RQ1. Does the most common websites that use some kind of reputation system have a high level of usability?**

#### **5.1.1 Within heuristic evaluation**

The websites that has been evaluated are among the most common websites in the world and they need to have a high level of usability, but unfortunately there are some problems found in most of the assessed websites. The main problem was the confusion of the presented information, because they are presented in a similar way, so the user sometimes gets confused where the item details are and where the users review and other information are. This will make the user lose his focus on the needed information, so the process of the decision-making will be a very difficult task to the user.

There is another main problem is the quality of the reviews, most of the systems don't allow their users to vote on other review's usefulness even if these reviews were seen as questionable. The quality of the reviews should be managed in a better way as we know that many users depend on the reviews or comments to make decisions online.

We all know the relation between icons or links, and the functions they carry out are very important for the usability of the reputation system. We found from our evaluating for the chosen websites that number of reputation systems lack the consistency and standards and this will have impact on users' satisfaction with the system. So our recommendation is; the choice of icons or links should therefore be based upon similar use in other relevant systems which the user is likely to have used it before.

One of the recommendations to the reputation system is; it should be placed an incentives for a quick assessments, because the faster the user hands in his assessment, the faster a seller's misconduct can be discovered. Moreover, there should be no time restrictions for leaving an assessment, because many sellers use some tactics to prevent the buyers from writing negative assessment until the time restriction is over.

Most of the assessed websites lack a good help system, so to improve it; the websites reputation system should provide a clear and sufficient help and guidance to the users. There should be quick help and user friendly and sufficient information to support users when sometimes users end up in the wrong place or can't find what they're looking for, e.g. dispatch rate or payment method. The system should provide them with other useful options, like other help topics, a forum and so on via relevant hyperlinks. The help system should at least be:

- Easy to search for solutions to specific problems.
- Help content must be clear and illustrative, and can be read quickly.
- Each help section is as self-contained as possible; to prevent confusion, help systems with a lot of content should be carefully structured by topic.

### 5.1.2 **Within think aloud testing**

Depending on the test participant point of view most of the participants were agreed that most of the evaluated websites lacks high level of usability. Most of them faced many problems when they used the websites. The main problem was that most of the participants were annoyed that they didn't understand exactly how the star rating is calculated and what it refers to. They thought that the star rating in most of the evaluated reputation system is too general and they can't figure out what are the attributes or the parameters that have been taken into account to build this rating. During the study periods we tried to search and contact many these websites to know how they generate their reputation score, but unfortunately they were unwilling to discuss with us how their reputation system is working, to ensure their generated reputation score is a true score, and could not been manipulated. So our recommendation is; these websites should be more open about the way their reputation system works. This would help the users to understand and trust the reputation scores and then they will use it.

Another recommendation it should be added a detailed feedback about the seller or the business. A differentiation in behavior of the seller e.g. product quality, postage and packaging and communications, and a differentiation in behavior of the business e.g.

product quality, price and service allows a better evaluation of a seller and business respectively. This would also bring clearness to the users, and will allow them to take a fair decision (separate assessment for each reputation attributes) instead of place an entirely negative assessment, which most users want to avoid. Hence, this feature might lead to a more honest rating behavior and richer information on the one hand and on the other hand will help the seller/business to discover what they're good at and what they can do better to score higher rating.

Another main problem that faced the participants was the difficulty to categorize the reviews, especially to display the negative reviews. So our recommendation it should be a possibility to quickly and easily filter the negative reviews, because this feature will help the users to get a complete picture of the seller and his activities. So the sites should add the "rating" criteria to their reviews sort, and there should be possibility to display the reviews in different period of time e.g. last 3 months.

There are other problems that faced the test participants like navigation and find enough information about the provider or the product. There should be enough information about the seller and products e.g. comparison the reviews about a product with other similar products, so the consumers can decide which product will choose according to his needs.

## **5.2 RQ2. Does the reputation information representation have an impact on users?**

### **5.2.1 Within survey and interviews**

We found from our study that most of the users consider the most important reputation element affecting making decision are the product picture and description, number of positive feedback percentage and the seller rating stars. So we can find that the visual rating is very important to the users in their decision making due to its highly visible appearance.

So the importance of visual appearance recommends that the important reputation elements should be presented in a visually eye-catching way. So it should be an additional visualization of detailed profile, for example a bar chart visualization (like the one used by Yelp) can assist the user in quickly understanding the development of a user's reputation over time. And the reputation elements should be presented separately allowing the users to make their decision (conclusion), because some of the elements do not appear to be utilized such as the communication rating.

Another important point is the user profiles should also be presented in a better way, although most of the websites have some form of member profile feature, a set of pages where one can find out about individual users. User profile pages provide a space where users can write about themselves, and their interests. In addition, profile pages often include statistics about a user's activity, recommendations from other users and sometimes a rating that reflects his achievement in that community. The user profile should be presented in an attractive and motivational way in order to help other members to use this information to make a decision in engaging in a transaction or collaboration with him, as it is found in [20] the visualization would motivate the subjects to contribute more papers and ratings and to participate more actively in an online community by logging in more frequently and reading more papers.

We found from the study that it is clear that most of the users tend to read the reviews (e.g. feedback) before they make a decision regarding buying a product or doing business with a seller. So the visual and textual formats for reputation representation are both important and should be available to help users in their decision making.

The websites should help the users in making their decision, so for example if the goal of the system is to help the users decide whether to trust the seller or not, then the system should present clearly the seller's percentage of the completed transactions while the other elements for example his history as buyer is less important.

## **5.3 RQ3. Does the usability of the reputation system have an impact on the user review and decision making?**

### **5.3.1 Within survey and interviews**

From the conducted interviews and survey we found that the usability of the reputation system has a big impact on people's utility of the reputation system whether by making decision or leaving feedback.

We found that most of the users don't leave negative feedback, because they are afraid of getting a revenge rating conversely. They don't know in Amazon that Seller-to-buyer feedback is not counted toward a feedback average. Only buyer-to-seller ratings are factored into an average. And eBay bans sellers from leaving negative or natural feedback for buyers. So the reputation systems for these sites should make it obvious to their members about this important policy, because this feature has so many positives such as:

- The buyers will be more honest when they leave feedback since he will not be fear of a negative revenge feedback.
- The buyers will bid more since their trust in the feedback system and the sellers will increase.
- The good sellers will differentiate themselves from other sellers.
- The sellers will leave feedback upon payment more often in order to increase their chance of receiving positive feedback.

From the survey and interviews we found that most of the users will read the textual reviews of the seller/product before they will deal with the seller or buy the product. We found also that the higher the usability of the website is, the higher the users trust towards the website, and also their contributions to these sites whether leaving feedback or their future intentions to use the same websites to do business than other websites. Most of the users preferred to leave feedback in eBay and Amazon than to leave a review in Epinions, because they thought the review system in Epinions is very complicated and to write a review in it will require a long time.

One of the findings was the usability of the website is more important than other factors such as the reputation of the site. We found that the most important things that users want in a system before they can use it are:

- The users want to complete any task in an efficient and easy way, and they want to customize the system to behave the way they desire.
- The experienced users want more advanced functions or features that would enable them to be more productive.
- The users want the chances for their errors are minimized, and the system should help them to understand the problem when errors occur, besides the system should help them to recover from errors.
- The users want to navigate the website easily, i.e. they can find what they want easily.

# Chapter 6

## 6. Conclusion

---

The main purpose of our study is to evaluate the usability and to get an in-depth understanding of the reputation-based websites, and to analyze the influence of the usability of the reputation system on its utility; therefore we choose seven known websites that use some kind of reputation system. All the chosen websites is known and leader in their fields. From the previous discussion, it can be stated that the application of the usability evaluation methods and usability study as a whole was successful in identifying such usability problems. On the other hand, even if the Nielsen ten heuristics was applied by just one evaluator, it was possible to identify a range of usability problems, because we used think aloud test, which help us to find the usability problems from user's point of view.

During our study we worked out that the level of the usability needs more improvement. It is also observed that the usability of the reputation system has influence on the its utility. One of the important concluded points is: it should be clear to the user how reputation ratings are obtained and what they mean, because to help the user to trust and use these information, it is important that the user understands the way information is measured.

The reputation systems should be designed in such a way that allow the user to easily understand how reputation is measured (e.g. which factors are taken into account and their weightings), what implications lie behind a given reputation rating, how reputation is verified, in order to let the user use the reputation system.

We worked out also that users should be offered qualitative and quantitative assessment of reputation indications. When it comes to trust the users prefer the qualitative information, while quantitative indications may offer advantage for a quick evaluation. Hence using combination of quantitative and qualitative approaches is recommended, wherever the system allows it.

We can conclude that the usability and utility are inseparable, if the usability goals are not achieved in the website, the resulting product will be useless to the users.



# Chapter 7

## 7. Bibliography

---

1. Jøsang, R. Ismail, C. Boyd (2007): A Survey of Trust and Reputation Systems for Online Service Provision, Information Security Research centre, Queensland University of Technology, Brisbane, Australia, Preprint of article published in Decision Support Systems, 43(2) 2007, p.618-644.
2. J. Nielsen (1993), Usability Engineering, Morgan Kaufmann, San Francisco, CA, USA.
3. Family Health International, Qualitative Research Methods: A Data Collector's Field Guide, Module 1: Qualitative Research Methods Overview.
4. R. Molich (2007), Usable Web Design, 1. Edition, chapter 11, Nyt Teknisk Forlag.
5. P. Resnick, R. Zeckhauser (2001), Trust Among strangers in Internet Transactions: Empirical Analysis of eBay's Reputation system
6. Ebay.ca, <http://pages.ebay.ca/help/feedback/allaboutfeedback.html> (Accessed 7 July 2012)
7. Amazon.co.uk, [http://www.amazon.co.uk/gp/help/customer/display.html/ref=gp\\_m\\_b\\_h\\_e?ie=UTF8&nodeId=471044](http://www.amazon.co.uk/gp/help/customer/display.html/ref=gp_m_b_h_e?ie=UTF8&nodeId=471044) (Accessed 8 July 2012)
8. Yelp.com, <http://www.yelp.com/faq> (Accessed 9 July 2012)
9. Trustpilot.com, <http://www.trustpilot.com/faq> (Accessed 9 July 2012)
10. Epinions.com, [http://www.epinions.com/help/faq/?show=faq\\_recognition](http://www.epinions.com/help/faq/?show=faq_recognition) (Accessed 10 July 2012)
11. iTunes Application.
12. Youtube.com, <http://support.google.com/youtube/?hl=en> (Accessed 15 July 2012)

13. Y. Shen (2010), the Simulation of Reputation Systems in Widget Sharing Communities: A Comparison Study, University of Helsinki, Finland.
14. S. Ruohomaa and L. Kutvonen (2005), Trust Management Survey, University of Helsinki, Finland.
15. X. Faulkner (2000), Usability Engineering, Palgrave, New York, N.Y., USA.
16. J. Nielsen, Ten Usability Heuristics:
17. [http://www.useit.com/papers/heuristic/heuristic\\_list.html](http://www.useit.com/papers/heuristic/heuristic_list.html) (Accessed 9 April 2012)
18. E. Carrara, G. Hogben, Enisa (Oct. 2007), Reputation-based Systems: a security analysis, ENISA Position Paper No. 2, European Network and Information Security Agency.
19. P. Resnick, R. Zeckhauser, J. Swanson, K. Lockwood (March 2004), The value of reputation on eBay: a controlled experiment.
20. J. Nielsen, Introduction to Usability: <http://www.useit.com/alertbox/20030825.html> (Accessed 7 April 2012)
21. J. Vassileva, L. Sun (2007), An Improved Design and a Case Study of a Social Visualization Encouraging Participation in Online Communities, this work is supported by the NSERC Discovery Grant Program, Canada.
22. J. Rubin, D. Chisnell (2008), Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests, second edition, Wiley Publishing, Inc., Indianapolis, Indiana, IN, U.S.A.
23. J. Nielsen, Thinking Aloud: The #1 Usability Tool: <http://www.useit.com/alertbox/thinking-aloud-tests.html> (Accessed 16 April 2012)

# Appendix

---

## Appendix A. Usability Test Script

### 1. Test facilitator:

**Test facilitator is neutral** – I have not participated in the development of the websites. Please provide as many comments as you possibly can.

**This is not an examination** – in any case not of you. You can do nothing wrong during this test. If you have any problems with the website, it is the website's fault and not yours.

**I will answer only very few questions from you** – Sorry! But please ask questions anyway. The questions may give me important clues. I may ask questions you will find odd. The reason is that we want to understand how you interpret this website.

**I will ask you to think aloud** during the test. This will enable me to follow your thoughts. I am interested in any and all comments; Positive as well as negative.

### 2. Pre-test interview questions

1. Have you ever visit these sites before? If so, what was your experience?
2. Have you ever visited websites for similar websites? If so, which websites? What was your experience?
3. What are your expectations to a based reputation website?
4. What is your level of familiarity with using the Internet?
5. What facilities would you want on a site of this type?
6. Do you have any questions?

### 3. Test tasks:

1. What comes first to your mind when you want to buy/reserve something? Is the reputation of the seller or the facility? Or the price? Or location? (all)  
(Test of what if the seller reputation attracts buyer).

2. You would like to know how the reputation system is calculated to the sellers. (all)  
(Test if there is a description of how the calculation of the reputation is done).
3. Considering the home page of this site, do you understand clearly how the reputation system of this site works? (all)  
(Test if the user understands the representation of the reputation).
4. Considering the home page of this site, you want to read the negative reviews about any (seller/business). (all)  
(Test if the user can sort the reviews by "rating").
5. You want to go to a restaurant find a restaurant that has a very good reputation. (YELP)  
(Test of what attracts most the customers according to the usability representation).
6. You want to go buy an iPad 3 find an iPad that has a very good review from trusted authors. (Epinion)  
(Test if the user understands the member status).
7. You are convinced with f.ex. Call me mobile company and you want to leave them a feedback to give them your opinion in TRUSTPILOT site. (TrustPilot)  
(Test if leaving feedback is easy or difficult for the users).
8. You want to install an app about messenger into your iPhone, which one you will choose? (ITunes. Apple)  
(Test of what attracts most the customers according to the usability representation).
9. You want to see which song for Adele that has highest rating. (YouTube).  
(Test if the site has an obvious reputation system).
10. You have now received the item you have bought, but it turns out that there is a fault in the item. What do you do? (eBay & Amazon)  
(Test if the user will leave a feedback?).
11. You want to buy a battery for your laptop from a trusted seller? (eBay & Amazon)  
(Test of the website's ability to provide enough information about the sellers).
12. You are neither convinced nor upset with your item, and you want to leave a feedback. (eBay / Amazon)

(Test if the user understands the reputation system).

13. You have forgotten to leave a feedback to a seller, and now after 3 weeks you remember that, what shall you do? (eBay & Amazon)  
(Test if the user knows the rules of the reputation system).

#### **4. Interview after test**

1. Do you find that the content is relevant for you?
2. Did the website meet your expectations?
3. Were the tasks realistic?
4. What tasks were missing?
5. If the test participant has visited other websites for other similar websites: How does this website compare?
6. Do you have any comments regarding the graphics on the website?
7. What are the two best things about the website?
8. What are the two things that most need improvement?
9. Will you use this website in the future?
10. How did you find this system to you? (very pleasing, pleasing, average, irritating, very irritating)
11. Would you recommend this website to a friend?