Empirical studies of Quality of Experience (QoE) – A Systematic Literature Survey

Kestutis Mintauckis

Master of Science Thesis

UNIVERSITY OF OSLO

30 April 2010

Empirical studies of Quality of Experience (QoE) – A Systematic Literature Survey

Kestutis Mintauckis

April 2010

Department of Informatics at the University of Oslo
Master of Science Thesis in computer science
60 credit points

Printed in Norway.

Contents

List of Tables

List of Figures

Abstract

Quality of Experience (QoE) is a relatively new phenomenon. The main focus of this thesis has been to conduct a systematic literature survey of research done in the field of QoE over a ten year period. The method, developed by A. Fink, has been used to survey empirical studies. A framework of QoE has been developed, which created the possibility of grouping together and analysing all the studies in a common framework.

In total, 44 studies were analysed. 66 per cent of them were studies with human participants and 34 per cent of them were studies without human participants. The majority of the selected empirical studies have analysed the sub-aspect ‘satisfaction’. Among other vital sub-aspects, which were of interest to researches, were ‘usefulness’, ‘ease of use’, ‘communication’, ‘loss/packet loss’, ‘delay’, ‘bandwidth’, and ‘jitter’. The results of this survey show that different sub-aspects depend on different services. It is not enough that one sub-aspect functions very well, because most of sub-aspects are closely related to each other. Therefore, it is very important that sub-aspects, which are dependent on each other, are functioning as one group to achieve higher QoE on user experience.

This thesis may contribute to deeper understanding of the phenomenon QoE. Knowledge of QoE can bring in new ideas and new possibilities for developing a new system or products for achieving satisfaction of user experience.

Key words: Empirical studies, Quality of Experience (QoE), Quality of Service (QoS), literature survey, aspects, sub-aspects, ISO standards, models, interaction design criteria, framework, empirical strategies, methods.

Acknowledgements

This master's thesis has been dependent on the vast knowledge and enthusiastic collaboration of various people and I would like to show my gratitude to them.

First of all, my supervisor Amela Karahasanovic deserves my deepest gratitude for the professional support and inspiration she has provided for me. As a mentor, she has created an environment in which my freedom to pursue new topics and ideas with curiosity has combined with her sharing of her invaluable expertise and research insight with me. Without her help, it would have been difficult to complete this work.

The people who work at the library in the Department for Informatics have also contributed to my present thesis. Especially, I would like to thank to Knut Hegna and his colleagues for assisting me with different questions and for all their help and good service.

I am grateful to SINTEF and the Department of Informatics, University of Oslo for providing a good working environment.

I would also like to thank Arne Bakstad and colleagues from my work at Oslo Universitetssykehus for providing enough time and freedom to work on this thesis.

I would also like to thank my good friend and fellow student Trond Sorvoja for useful discussions on my thesis and a lot of evenings together with me at the University reading room. I am also thankful to my friend Anders Gjørven for his encouragement

I must mention my gratitude to my life companion Ramune Grambaite, who has given me the necessary time and unforgettable support to as well as valuable language consultations during all stages of the writing process.

Barry Kavanagh has proof-read my thesis. Thanks!

Last but not least, to my family, loved ones and friends for their support throughout this period.

1. Introduction

1.1 Motivation

> *Science is built up of facts as a house is built of stones, but an accumulation of facts is no more a science than a heap of stones is a house.*
>
> **Jules Henri Poincaré (1854 – 1912)**

The concept of quality is complex. To understand what quality is, users might want to specify more precisely the context in which they are discussing it, for example, the quality of clothes, sound quality, video quality, or software quality in organizations etc. There are plenty of companies and universities using IT systems that enable them to perform high quality tasks to satisfy their customers, otherwise the IT systems would be useless. This indicates specifically that those who develop software or IT systems should obviously understand the phenomenon of quality, which is very important in computer science. The developer, or others involved in this process, should have a clear definition of quality, how the users can measure and evaluate it and how quality is defined and characterized.

There are still a lot of discussions about the pros and cons of this phenomenon. A few researchers, e.g., McCall et al. (1977) and Boehm et al. (1976) have carefully analysed, defined and described this term. Another significant source in this field, providing a clear definition, is ISO standards i.e. ISO 9241-11, ISO 9126, J.Nielsens 'A model of system acceptability' has defined it clearly. Since understanding of this phenomenon is not mature, more research in this area is needed.

Development in the future might involve the integration of real-time multimedia services by service providers. Triple-play (IPTV, IP Telephony and Internet) services offer a data stream such as voice over IP (VoIP) and a video stream at the same time over an IP network. Voice over IP will integrate the plain switched telephone network while Internet Protocol Television (IPTV) will integrate the delivery of television broadcasts to Internet end-users into IP networks. Providers of the services have a challenging task with this integration process and a sufficient amount of quality of service must be provided in best-effort IP networks to achieve high customer satisfaction. This shows again that quality is a significant phenomenon in the context of information technologies.

Another important concept and related to quality is the phenomenon 'Quality of Experience' (QoE). To understand this concept, the user might have a clear definition, how to measure it and how to classify it, but there are still very few information sources about QoE. This present thesis aims to find out how QoE has been defined in different articles, how many aspects this phenomenon has and how QoE affects user experience.

The main focus of this thesis is to provide an overview of research done in the field of QoE. The studies, which have described, analysed and evaluated the phenomenon of QoE, have been included in the literature survey. This study aimes to find out as much as possible about QoE

generally or QoE from the user's perspective, i.e. how new technology can help users in their every day lives or how easily or quickly users will get used to a new product. QoE, in other words, is related to user experience.

Additionally, in order to be able to perform this survey, I aimed to develop a framework that can help to classify and evaluate all aspects used in the selected empirical studies and to attain more knowledge about all aspects and sub-aspects that belong to QoE, i.e. 'Learnability and Usability', 'Appropriate', 'Manageable', etc. The scientific field of QoE is not mature yet, therefore it is very important to accumulate more details about aspects and sub-aspects that have been described by different researchers.

QoE is the main focus of the present thesis and there are many definitions of it. I provide one here:

DEFINITION -*'Quality of Experience (QoE or QoX) is a measure of the overall level of customer satisfaction with a vendor. QoE is related to but differs from Quality of Service (QoS), which embodies the notion that hardware and software characteristics can be measured, improved and perhaps guaranteed. In contrast, QoE expresses user satisfaction both objectively and subjectively. The QoE paradigm can be applied to any consumer-related business or service. It is often used in information technology (IT) and customer electronics' (Unknown, 2009b).* In addition, the QoE depends on the users, as some of them may be easier to please than others. QoE is furthermore affected by some major factors such as cost, reliability, efficiency, privacy, security, interface, user-friendliness and user confidence.

1.2 Objective

The objective of this thesis is to increase our understanding of QoE by conducting a literature survey of 'Quality of Experience'.

To approach the objective, I had to address the following sub-goals:

- To provide an overview of the relevant definitions, models, ISO standards and empirical methods in relation to QoE.
- To develop a framework for evaluating different aspects and sub-aspects of QoE in selected empirical studies.
- To perform a systematic literature survey.

The theoretical background of 'Quality of Experience' will be described in detail later in this thesis (chapter 2.2). One of my main tasks has been to search different electronic literature databases and pick up relevant and meaningful articles for my thesis.

1.3 Research method

The research method used in this thesis is a systematic literature review. 'The need for a systematic review arises from the requirement of researchers to summarise all existing information about some phenomenon in a thorough and unbiased manner. This may be in order to draw more general conclusion about some phenomenon than is possible from individual studies, or as a prelude to further research activities' (Kitchenham, 2004). Some elements of Kitchenham's guidelines that were applicable were fundamentally maintained, although some modifications were necessary.

I decided to search the ten year period from 1999 until July 2009. I searched seven different electronic databases (see Appendix C). Then I went through each article, reading both titles and abstracts. Following this, I had approximately two weeks' break and then I went through the same results list one more time (see Appendix A, Figure 19 and Figure 20). Due to the capacity and time, I decided to include neither master's theses nor PhD dissertations.

In total, there are a few hundred articles found using different databases such as IEEE, ACM, and Google Scholar etc. A total of 44 articles about quality of experience were thoroughly analysed. Useful information about each study was collected on a separate list and stored in an Excel spreadsheet.

The reader of this thesis will note that the references section has been split into two parts:

- The first list, which I have used in the writing of the theory part of this thesis (see References A)
- The second list containing the studies I have surveyed (see References B)

1.4 Research context

This master's thesis has been done in the context of the *Road to media-aware user-Dependent self-aDaptive NETWORKS* (R2D2 Networks) project (Roldan, 2008), which is currently run by SINTEF. This project focuses on triple-play services. R2D2 Networks' main outcome is to be a user-dependent media-aware self-adaptive network that will help the network to learn what, when, where and how services and resources are used.

Nowadays, there are access networks that have been implemented in offices, schools, households, etc. Neither the access network operator nor for example the household has the power to control which services and what quality of services are or should be delivered through the access link. Services are delivered in the best possible way but in the case of bandwidth failures or link congestion, all services used by the abovementioned household members will be affected. The result will be long delays and decreased user experience or QoE.

With triple-play like IPTV, IP Telephony and internet, all these multimedia services are delivered in one single broadband access link. The advantage is that triple-play is in full operation, i.e. household members might be doing different things at the same time, some for example watching

several IPTV channels, downloading streaming media on a computer, executing a file-sharing application or using IP Telephony. The triple-play project is a three year project that was started in June, 2009 in Madrid, Spain.

By obtaining a clear overview and understanding of users and usage together with QoE monitoring tools, this project will modernize the current home and at the same time will help to build network architecture with a better use of the connectivity resources in relation to the Quality of Service provided.

R2D2 Networks' aim is to change the European Telecommunications industry from an infrastructure and connection driven industry to a service and content driven industry. Another thing is that broadband connectivity, which is called 'product' will no longer exist, but instead of will be replaced by 'personalized services'.

R2D2 Networks is an international project involving countries such as Spain, Norway, Sweden, France, Greece and Cyprus. There are two companies Uninett and SINTEF, which participate in this project in Norway.

QoE is an important concept for the R2D2 Networks project. To provide the triple-play service in the best possible way, or best-effort fashion, is the one of the most important tasks for the provider. By delivering good services, it is possible to achieve user experience satisfaction.

This thesis will present and analyse sub-aspects such as 'satisfaction', 'ease of use', 'delay', 'bandwidth', 'jitter', etc. These aspects are closely related to the phenomenon QoE and the phenomenon Quality of Service (QoS). Services such as IPTV or IP Telephony are delivered throught the internet. Therefore, QoS is a very important concept for R2D2 Networks too.

To find out as much as possible about the relatively new phenomenon QoE is the purpose of this thesis, and the findings may contribute to another project that is working in parallel, namely R2D2 Networks.

1.5 Contribution

The phenomenon of QoE is relatively new and researchers started to use it in their work just few years ago. My contribution is described below.

- For a better understanding of the concept of quality (Chapter 2.1), I decided to give some definitions of it. Also, ISO standards (Chapter 2.1.2), a model of system acceptability (Chapter 2.1.4) and interaction design criteria (Chapter 2.1.6) have been presented in this thesis in order to help the readers understand the concept. At the same time, an overview of empirical strategies (Chapter 2.4), involving an experiment, a survey and a case study, has been included. It was also possible to find out which strategies were most often used by researchers. Finally, some methods have been described (Chapter 3), which have been developed by well-known researchers. One method, which was developed by A.Fink, could be applied in the undertaking of a systematic literature survey.

- By connecting together ISO standards, method and interaction design criteria, it became possible to develop a framework (Chapter 4). All aspects, which were selected from all studies, could be analysed and grouped by using this framework (Appendix F, Figure 24).

- The main focus of this thesis has been to complete a systematic literature survey. This survey gives a brief overview of work that has been done over a ten year period by researchers working on the concept of Quality of Experience. All the selected studies were scrutinized for each of the categories *subject* (Chapter 4.1), *type of study* (Chapter 4.2), *aspects* (Chapter 4.3), *purpose* (Chapter 4.4) and *results* (Chapter4.5). This created the possibility to compare the studies with each other and make a conclusion from the findings. Finally, potential threats to internal, external and construct validity were described (Chapter 6). A brief overview of these threats may be helpful in understanding them and preventing them from biasing this thesis.

Furthermore, this thesis has contributed to new knowledge about all aspects and sub-aspects that belong to QoE, i.e., '*learnability and usability*', '*appropriate*', '*manageable*', '*aesthetic experience*' (see Figure 6), etc. The scientific field of QoE is not mature yet, therefore it is very important to get more details about aspects and sub-aspects that have been described and analysed by different researchers.

The findings from the present thesis could be useful for researchers, scholars, and practitioners. They could apply them in further research. Knowledge of QoE can bring in new ideas and new possibilities for developing a new system or products for achieving satisfaction of user experience.

1.6 Structure

The remainder of this thesis is structured as follows:

Chapter 2 *page 15*
Background
This chapter describes some definitions related to the concept of quality of experience in this thesis. It also presents ISO standards, a model of system acceptability, and interaction design criteria, to help the readers to understand the concept. Finally, some empirical strategies are discussed.

Chapter 3 *page 28*
Method
This chapter contains the description and discussion of different methods. The method developed by A. Fink, and applied in undertaking a systematic literature survey, is described in detail.

Chapter 4 *page 34*
Analyses
This is one of the main parts of the literature survey.
First of all, a framework was developed. Secondly, selected studies were analysed for each of the categories *subject*, *type of study*, *aspects*, *purpose* and *results*.

Chapter 5 *page 58*
Discussions and Summaries
Discussion of the results from selected studies is provided here.

Chapter 6 *page 64*
Threats to Validity
Definitions and descriptions of the internal, external and construct validity of the findings in the present thesis are given.

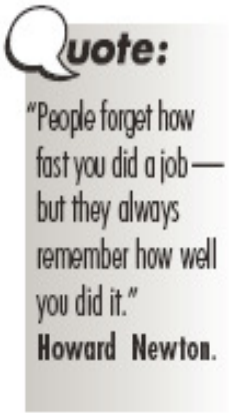
Chapter 7 *page 66*
Conclusion and Future work
The conclusions from the review of the related work have been extracted. At the end, some ideas for future work are provided.

page 69
List of Acronyms and Abbreviations
A list of acronyms and abbreviations is given.

page 71
References A
A list of references, which belong to the theory part (Chapter 2 and Chapter 3), is given.

page 76
References B
This part is a list of the 44 studies that have been selected and analysed for this thesis.

2. Background



Information and communication technologies (ICT) allow users to participate in a rapidly changing world. They include any communication device, such as a computer, networks, network hardware, satellite systems, other data processing and transmitting equipment, and software. This term can also be associated with various services and applications and communication between them, such as distance learning or videoconferencing, etc. Nowadays ICTs are discussed in a particular context, such as ICTs in education, health care and libraries (Unknown, 2009e; Yeh, 2009).

The ICT concept is broad and only a few selected areas have been analysed in this thesis, such as Software Quality (SQ), HCI and telecommunication. This chapter presents concepts and definitions related to different aspects of quality within ICT.

Firstl of all, a short definition of software quality is given. Secondly, a few ISO standards, which describe each different aspect of HCI usability in more detail, are presented. Thirdly, there is a description of qualities of user interface and ACM awards criteria.

2.1. Quality factors

On the topic of quality and evaluation, it is best to look at the most relevant definitions that provide a good explanation and deeper understanding, for example ISO standards, usability, Nielsen's model of system acceptability (Figure 2), the ACM/interaction design awards criteria (Figure 3), etc. The reason for that is as follows.

Quality experts state that most aspects of quality are more or less related to users' needs and these aspects are considered during definition and assessment. This conforms to ISO standards' definition of quality, 'the totality of characteristics of an entity that bear on its ability to satisfy stated and implied needs' (Kitchenham et al., 1996). Secondly, there is a close relationship between software quality, usability and interaction design related to ICT. Thirdly, usability is just one small part of interaction design. Therefore usability and interaction design are dependent on each other.

In order to take advantage of the discussions within the thesis and make understanding easier, the following definitions are given.

2.1.1. Software Quality

Software Quality (SQ) is a complex concept in computer science. Most people are more or less affected by the quality of software, because IT organizations, universities, etc. use different kinds

of software solutions for their customers. Each customer may have different meanings about software quality and yet still believe that quality is important and can be improved. Both large and small companies invest much time, a large amount of money and expend much effort in improving software quality (Kitchenham & Pfleeger, 1996). Some companies focus on the process, while other companies adopt a more product-based approach to achieve software quality – and these two strategies have led to the Malcolm Baldridge Awards (Unknown, 2009a).

Garvin (1984) has stated that 'quality is a complex and multifaceted concept' and that the concept of 'quality' has five different perspectives. Kitchenham et. al.(1996) have also written about software quality with the view that it can be described from five different perspectives.

1. ***Transcendent definition:***
   The transcendent view defines quality as something that can be recognized but not defined.

2. ***User definition:***
   The user view defines quality as fitness for use.

3. ***Manufacturing definition:***
   The manufacturing view defines quality as conformance to requirements, design, or specification.

4. ***Product definition:***
   The product view sees quality as tied to the inherent characteristics of the product.

5. ***Value definition:***
   The value-based view defines quality as dependent on the amount a customer is willing to pay for it.

Other researchers claim that quality is a value, and it has been described for example by Weinberg (1991) as follows: 'Quality is value to some person'. This definition stresses that quality is inherently subjective. It means that people will understand the quality of the same software in different ways (Renier, 2009).

In (Berander et al., 2005), chapter 1, 'Software Quality Models and Philosophies', two major points about software quality are made.

1) **Conformance to specification**: quality that is defined as a matter of products and services whose measurable characteristics satisfy a fixed specification – that is, conformance to a specification defined beforehand.

2) **Meeting customer needs**: quality that is identified independent of any measurable characteristics. That is, quality defined as the products or services capability to meet customer expectations – whether explicit or not.

The first point, 'Conformance to specification', is clearly described by Philip Crosby in his book *Quality is free: the art of making quality certain* (Crosby, 1979), while Shewhart presents

(Shewhart, 1931) quality as having both an objective and subjective side. Shewhart's explanation and description can be rendered as 'conformance to specification' and 'meeting customer needs'.

It has been explained above through a short description how quality can be defined, but it is also worth being able to measure quality. Software Quality (SQ) measures how well software is designed and how well the software conforms to that design (Renier, 2009).

2.1.2. ISO 9126 Standard

ISO 9126 is the software quality evaluation standard, which was adopted by the International Organization for Standardization (ISO) and the International Electro technical Commission (IEC) (ISO, 2009; McInnes, 2009) . In the first part of the standard, ISO 9126-1 established the quality model, which classifies software quality in a structured set. Six characteristics of this model are as following:

1. *Functionality* is the set of attributes that bear on the existence of a set of functions and their specified properties. The functions are those that satisfy stated or implied needs.
2. *Reliability* is the set of attributes that bear on the capability of software to maintain its level of performance under stated conditions for a stated period of time.
3. *Usability* is the set of attributes that bear on the effort needed for use, and on the individual assessment of such use, by a stated or implied set of users.
4. *Efficiency* is the set of attributes that bear on the relationship between the level of performance of the software and the amount of resources used, under stated conditions.
5. *Maintainability* is the set of attributes that bear on the effort needed to make specified modifications.
6. *Portability* is the set of attributes that bear on the ability of software to be transferred from one environment.

Each characteristic in software quality model can also contain one or more sub-characteristics. *Functionality* has five different sub-characteristics, i.e accuracy, compliance, interoperability, security and suitability, while *Efficiency* has just two: time behaviour and resource behaviour. These characteristics and sub-characteristics are defined by ISO/IEC and the reader can get a full list at a glance, from Figure 1.

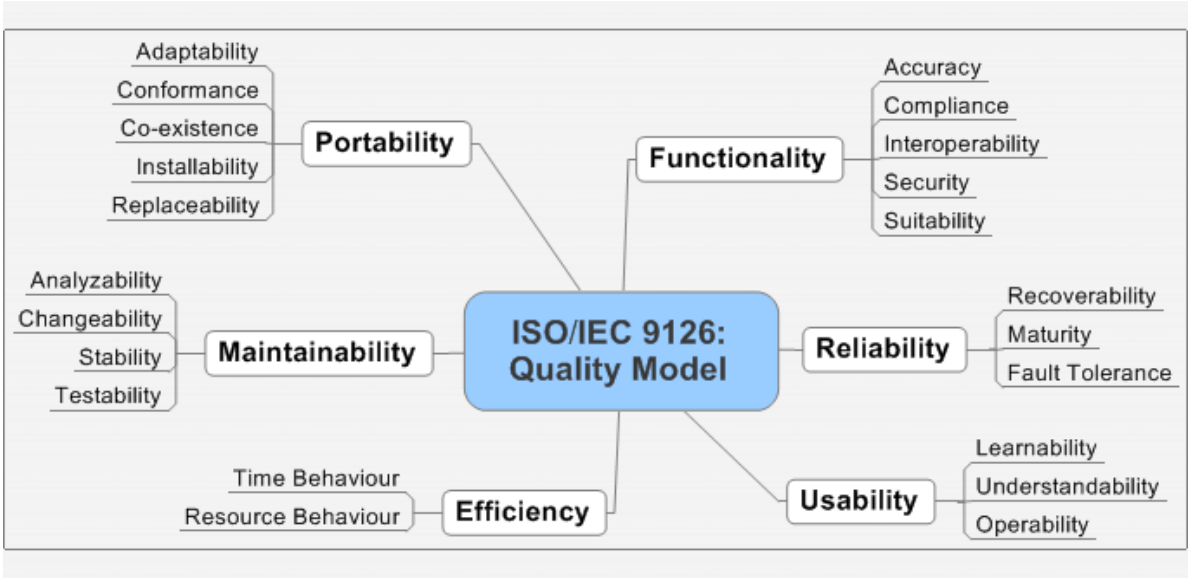


Figure 1: Software quality characteristics and sub-characteristics defined by ISO/IEC 9126.

It is important to follow two approaches in order to ensure product quality. The first one is assurance of the process by which a product is developed (i.e. ISO 9001, ISO 9000-3 provides guidelines for the application of the ISO 9001) and the second one is the evaluation of the quality of the end product (i.e. ISO 9126). These two approaches involve the presence of a system for managing quality (ISO, 2009).

2.1.3. ISO 9241-11: Guidance on Usability (1998)

ISO 9241-11 explains how to identify the information that is necessary to take into account when specifying or evaluating usability, in terms of measures of user performance and satisfaction (Abran et al., 2003).

Furthermore, this standard gives a deep explanation of how to measure user performance and satisfaction and how any component of a work system affects the quality of the whole work system (Bevan, 2009; Harning et al., 2003).

This standard consists of the three components of usability, which are defined as:

1. *Effectiveness* – the accuracy and completeness with which users achieve their goals.

2. *Efficiency* – the resources expended in relation to the accuracy and completeness with which users achieve goals.

3. *Satisfaction* – the comfort and acceptability of use.

According to standard ISO 9241, software is considered usable when the user can execute his tasks effectively, efficiently and at the same time user feel satisfaction in the specified context of use.

Another consideration, *satisfaction*, is important for the Graphical User Interface (GUI), according to Hassenzahl et al. (2001), who wrote 'The latest definition of usability is, it recently acquired a new associate, the so-called *joy of use*'. Success key is when users enjoyed a software product; it is likely that most of users are satisfied with this product.

2.1.4. A model of system acceptability

Jakob Nielsen's book *Usability Engineering* (Nielsen, 1993) gives a brief explanation of different attributes of the model of system acceptability. Figure 2 gives a more detailed overview of how attributes belong to each other.

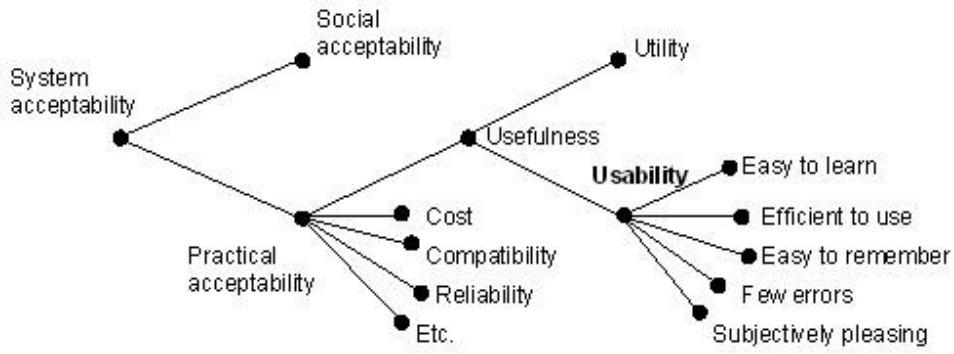


Figure 2: A model of the attributes of system acceptability (Nielsen, 1993, s.25).

The attribute *system acceptability* is possible to compare with *usability* to some extent, according to the author. System acceptability concerns the question of whether the system is good enough to satisfy all the needs and requirements of the system users or managers. Furthermore, this attribute can be split in two parts, *social acceptability* and *practical acceptability*. The latter covers various categories such as cost, support, reliability, compatibility and usefulness.

The attribute *usefulness* is concerned with whether the system can be used to achieve some desired goal, and this also can be split into two parts, *utility* and *usability*. The former is used to describe the extent to which the product provides the right kind of functionality to help users perform relevant tasks (Preece et al., 2006), while *usability* analyses the question of how well users can use that functionality (Nielsen, 1993).

Nielsen wrote that *usability* characterizes all aspects of a system with which a human being might interact, including installation and maintenance procedures. Furthermore, the author describes the following five main attributes of usability:

1. *Learnability* – the system should be easy to learn so that the user can rapidly start getting some work done with the system.

2. *Efficiency* – the system should be efficient to use, so that once the user has learned the system, a high level of productivity is possible.

3. *Memorability* – the system should be easy to remember, so that the 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.

4. *Errors* – the system should have a low error rate, so that users make few errors during the use of the system, and so that if they do make errors they can easily recover from them. Furthermore, catastrophic errors must not occur.

5. *Satisfaction* – the system should be pleasant to use, so that users are subjectively satisfied when using it; they like it.

2.1.5 Qualities of user interface

Interaction Design has six main qualities of successful user interface (Unknown, 2009f).

1. *Usefulness*: are users' needs satisfied by the interface functionality?

2. *Learnability*: how easy is it for the user to fulfil basic tasks when using the system for the first time?

3. *Efficiency*: after the user knows the interface, how fast is s/he able to accomplish the given tasks?

4. *Ease of memorization*: when the user returns to the interface after a while, how easily does s/he find the various functions again?

5. *Reliability*: is the interface conceived in such a way that the user makes as few mistakes as possible?

6. *User-friendliness*: does the user like using the interface?

After exploring the Internet, it may be concluded that many different definitions of 'Interaction Design' have been used in the literature. There has been some focus on the experience of using computers and the following definition has been given:

'Interaction design concept is used to describe how people interact with computers and communications. Interaction design also reveals the value of a communication service to its users, and their quality of experience at the time of using it' (Thackara, 2001).

There are nine criteria under which interaction design can be defined as successful and these are described in the next section.

2.1.6 Criteria for the quality of interactive products

When all the aspects of how people use an interactive product are satisfying or engaging, the way it feels to have a new product in their hands, if they understand quickly the product functionality, how they feel about it while they are using it etc., then this is called 'quality of experience' (Alben, 1996). Good and great design is complex and not easy to define. Criteria have been identified for effective interaction called 'the ACM / interaction design awards criteria' as shown in Figure 3. Each criterion of effective interaction is explained by Alben et al.(2003).

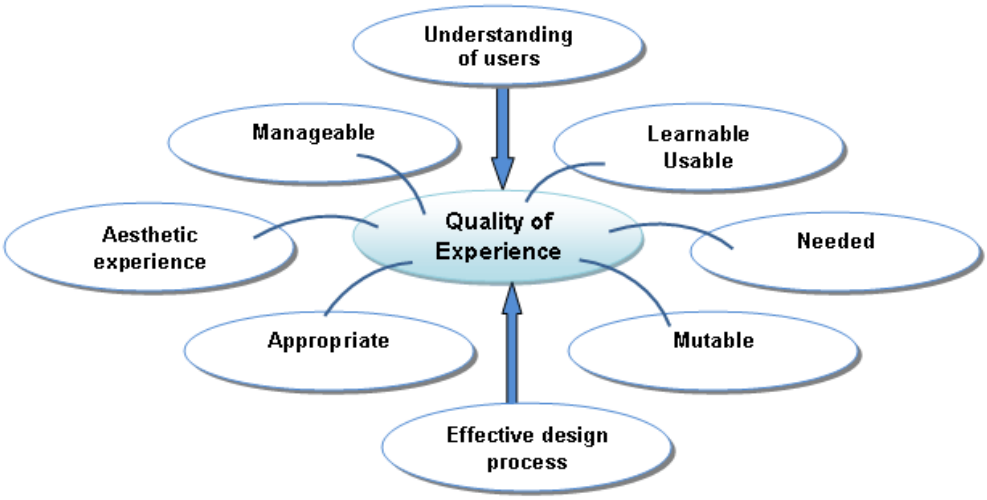


Figure 3: The ACM/interaction design awards criteria (Alben et al., 2003).

QoE is based on several criteria, which are listed below and explained by the authors.

1. ***Quality of Experience*:** all these criteria together lead to one question, does effective interaction design give users a successful and satisfying experience?

2. ***Understanding of users*:** does the team that designed the product understand the task, the environment of the user, and how is learning reflected in the product?

3. ***Effective design process*:** effective design process describes how well-thought out and well-executed the design process should be in order to achieve results. Which methods were used to solve the major design issues? Did designers use methodologies such as for example user involvement, iterative design cycles, and interdisciplinary collaboration?

How well were practical issues such as for instance budgeting, scheduling, and interpersonal communications managed?

4. ***Needed*:** this criterion concentrates on needs and their satisfaction, in order to result in a successful project. The product should make a significant social, economic or environmental contribution.

5. ***Learnable and usable*:** this criterion describes how easy it is to learn and use the product for the customer/user. How much experience and which skills does the user need for solving the problems? How easy is it for the user to proceed with the task using the product from the start to completion?

6. ***Appropriate*:** this criterion describes how the design of the product solves the problem at different levels and if the product serves customers in efficient and practical ways.

7. ***Aesthetic experience*:** aesthetic experience explains if the product is conceived as aesthetically pleasing and sensually satisfying. Moreover, this criterion describes whether the product is cohesively designed and has excellence across graphic, interaction and industrial design. There is also an investigation into whether the product performs well within technological constraints.

8. ***Mutable*:** this criterion describes the designer's consideration of whether mutability is appropriate or not. Another consideration is how well the product might be adapted to suit particular needs.

9. ***Manageable*:** this criterion describes whether the design of the product supports the entire context of use. For example, are the needs such as installation, training and maintenance, considered in both an individual and an organizational sense?

2.2 Quality of Experience and Quality of Service

This chapter presents some definitions. Firstly, a description of the phenomenon of quality of experience (Chapter 2.2.1) is given. Secondly, a description of quality of service (QoS) (Chapter 2.2.2) is given. Finally, there is a short presentation of QoE versus QoS (Chapter 2.2.3).

2.2.1 Quality of Experience

The term quality of experience (QoE) is rather new in multimedia contexts such as voice, video, IP Telephony, Internet Protocol Television (IPTV), etc. Different definitions of quality of experience have been used in the literature. Patric et al. (2004) defined this term as 'the characteristics of the sensations, perceptions, and opinions of people as they interact with their environments. These characteristics can be pleasing and enjoyable, or displeasing and frustrating.

Many factors contribute to a user's QoE of particular device or piece of software, including its appropriateness, effectiveness, learnability, and reliability'.

Goodchild (2005) states that 'Quality of Experience is the overall performance of a system from the point of view of the users. QoE is a measure of an end-to-end performance levels at the user perspective and an indicator of how well this system meets the user needs'.

Siller et al. (2003), state that 'the user's perceived experience of what is being presented by the Application Layer, where the application layer acts as a user interface front-end that presents the overall result of the individual Quality of Services'.

Perception of QoE as experienced by the end-user is crucial to comprehending what a customer might really want. Therefore, it is possible to understand the relationship between QoE and QoS. The combination of the two gives a service provider the possibility to find out how to improve, progress and assure the quality of customer experience. User and network perspectives are complementary to each other and are combined in one integrated QoE/QoS layered model. A more detailed explanation can be found in Zapater et al. (2007).

The QoE refers to the end-user's experience with the delivered service. QoE can be measured directly through user tests. It is expressed in terms of the Mean-Opinion-Score (MOS) but there are other methods that can be found to determine QoE, such as E-model ITU-T Rec. G.107 (Appendix D), instrumental metrics for example PESQ, ITU-T Rec. P.862 (Appendix D), or Rubino's Pseudo-Subjective Quality Assessment (PSQA) method (Rubino, 2005), Appendix D.

Research on audio QoE has matured, but research on video QoE is rather new and there is still a lack of a comprehensive metric. Moorsel et al. (2001) discussed how internet services have become increasingly popular. This may be due to three types of service models, such as business to consumer services, business to business services and the service utility model using service providers. Siller et al. (2003) explained that QoE is evaluated using weighted factors given by network QoS (NQoS) metrics such as jitter, delay and packet loss. Moreover, the term QoE is rather new in a technological context and it finds common use in company white papers that are related to systems and network services. Furthermore, the QoE is important for further Next Generation Internet (NGI) services, (O'Neil, 2002).

2.2.2 Quality of Service

Quality of Service (QoS) is a set of technologies for managing network traffic, and these create the possibility of reducing costs in effective ways so that the user experience can be enhanced, whether in a home or an enterprise environment. Through using these QoS technologies, the measuring of bandwidth and the detection of changing network conditions (such as availability or congestion of bandwidth) are enabled. It is also possible to prioritize throttle traffic such as traffic for latency-sensitive applications, i.e. voice or video (Unknown, 2009c).

DEFINITION - 'On the Internet and in other networks, Quality of Service (QoS) is the idea that transmission rates, error rates, and other characteristics can be measured, improved, and to some

extent, guaranteed in advance. QoS is of particular concern for the continuous transmission of high-bandwidth video and multimedia information. Transmitting this kind of content dependably is difficult in public networks using ordinary 'best effort' protocols' (Ziegler, 2004).

The QoS mechanisms can be considered as two layers: the application layer and the network layer. The application layer consists of the parameters (i.e. resolution, frame rate, color, video/audio codecs) related to video application and these parameters might be managed to achieve expected QoE levels. The network layer consists of the parameters (i.e. jitter, delay, packet loss) that are related to the transport and network performance. In order to meet the adequate QoE level, these parameters are managed (Zapater et al., 2007).

Various services are provided by the application layer to achieve the required QoS and the network services performed by devices, for example switchers, routers. On the top of these layers, there might be a perceptual pseudo-layer. This layer is concerned with the end-user experience and it is an extension of the application layer defined in the OSI Model. A metric, which is used for measuring the performance of this perceptual layer, is Quality of Experience (Aldrich et al., 2000; O'Neil, 2002).

QoS includes metrics to measure and assure a certain integrated services, multiprotocol label switching (MPLS), traffic engineering, and constraint-based routing (Xiao et al., 1999).

2.2.3 QoE versus QoS

Goodchild (2005) explains the main differences between QoS and QoE, in that QoE depends on the user's perspective, while QoS depends on the network perspective.

Definitions from different authors show that QoE when compared to QoS is on a higher, abstract layer, which is considered to be a perceptual pseudo-layer (Siller et al., 2003). This pseudo-layer is more or less concerned with the end-user's experience and it might be an extension to the Application layer, which is defined in the Open Systems Interconnection Reference Model (OSI Model) (Bauer et al., 2004). Furthermore, the QoE can be described as an extension of the traditional QoS in the sense that QoE provides information regarding the delivered services from an end-user point of view (Lopez et al., 2006).

Multimedia plays a very important role in both movies and computer-based products. Therefore QoE is a significant issue. When a producer releases a product into the marketplace, the end-users usually do not care what technology has been used to construct the product. Users normally care much more about which problems this products can solve or their experience while using it (Jain, 2004). Quality of Experience has serious implications for designers, so that they design effective multimedia approaches and systems. Usually, designers use quality of service (QoS) to represent the design parameters of a product. Therefore, it may be questioned if QoS may be extended to QoE and how many and what factors should designers consider in developing measures for QoE. It is not easy to present these questions to experienced people, because experience is subjective. Jain writes, 'We all know that cultural background, socioeconomic status, and personal experiences shape how a person reacts to a story (either in movie or book form) or a Web

interface' (Jain, 2004). At times it is enough to change one small detail, such as a colour in an interface, and the effect will be different on the end-users. Experience is also context-dependent. This means that the same person can have a different experience of the same book or movie depending on their understanding of the situation or experience. Therefore, it is quite difficult to identify factors that determine the context that influences experience.

2.3 Empirical studies

The aim of this research is to survey empirical studies on QoE. Therefore, brief descriptions of the main concepts are given here.

Empirical research is research that focuses on either direct or indirect observation as a test of reality. This kind of research may be conducted according to hypothetical-deductive procedures (Fisher, 2009).

'Empirical studies are used to investigate the effects of some input to the object under study. To control the study and to see the effects, we must be able to both measure the inputs in order to describe what causes the effect on the output, and to measure the output. Without measurements, it is not possible to have the desired control and therefore an empirical study cannot be conducted' (Wohlin et al., 2000).

There are two types of research: qualitative and quantitative, and they have different approaches to empirical studies. 'Qualitative research is concerned with discovering causes noticed by the subjects in the study, and understanding their view of the problem at hand. The subject is the person, which is taking part in an experiment in order to evaluate an object' – Wohlin et al. (2000). More details and discussions about relevant qualitative research have been described by Dittrich et al. (2007).

Quantitative research aim is to classify features, construct statistical models and explain what is observed. The researcher has clear opinion what he/she is looking for and might very carefully design all aspects of the study before data is collected. Quantitative data is more efficient and it is possible to test hypotheses. This type of research can also involve analysis of data such as pictures (i.e. audio or video) or words (i.e. interviews), etc. (Neill, 2007).

2.4 Overview of empirical strategies

Research strategies depend on the purpose of the evaluation. Different techniques, tools or methods have been used for empirical investigation. Because of this, it can be undertaken by a few different strategies, which are presented as follows (Wohlin et al., 2000).

2.4.1 Case study

A case study is an empirical inquiry that analyses a contemporary phenomenon, which will be investigated within a real-life context and the boundaries between real-life context and a contemporary phenomenon are not obvious (Yin, 2003). In other words, the case study method is used by a researcher who intentionally wants to cover contextual conditions. These conditions might be highly pertinent to the phenomenon the researcher is studying. In other words, the case study method may investigate a particular problem or situation in for example a company, and this investigation may be undertaken both directly, through interviews and indirectly by analysing documentation or reports that belong to the company (Creswell, 2003).

Some aspects of a case study inquiry are as follows.

- It analyses one distinctive situation, containing more variance of interest than data points.
- There are multiple sources of evidence.
- The prior development of a theoretical proposition has benefits and it will guide the data collections and analysis.

2.4.2 Experiment

An experiment is a method for investigating the relationship between different variables, or just to test a hypothesis. The experiment is done when the researcher can analyse or evaluate behaviour both directly, precisely and systematically. Laboratories are the right places to do this work and each experiment may focus on one or two isolated variables (Yin, 2003).

Experiments are used in both natural sciences and social sciences. This method can be also used to help solve practical problems and to support study. Also this method might be used to negate theoretical assumptions (Zekone, 2009).

2.4.3 Survey

A survey is a method that presents a quantitative description. The survey method can be very useful when a researcher wants to collect data that cannot be directly observed, for example people's attitudes towards one product or another, or users' opinion on social services, etc. (Czaja et al., 2005). Using these results, researcher can generalize or make claims about the populace. Researchers can also explain if the survey is *cross-sectional* or *longitudinal*. The former collects data at a single point in time, while the latter gathers data over a period of time. To collect data for a survey, there are some alternative methods such as on-line or self-administred questionnaires, structured record reviews, interviews or structured observations (Czaja & Blair, 2005; Fink, 1995).

2.4.4 A research literature review

To begin with, there should be a short explanation of the differences between a literature review and an academic research paper.

An academic research paper has as its main focus the support of an argument, while a literature review has to summarize and synthesize both ideas and arguments written by other researchers. An academic research paper has, generally, a limited number of sources, due to the emphasis on the argument. A literature review may have as well an 'argument' but it is not so important. The main focus is on covering a number of sources (UNC Press, 2009).

A research literature review is a systematic, explicit and reproducible method, which allows the researcher to identify, evaluate or synthesize recorded work produced by another researchers, scholars or practitioners (Fink, 2009 s.3).

Kitchenham has defined a systematic literature review (SLR) as follows: 'A systematic literature review is a means of identifying, evaluating and interpreting all available research relevant to a particular research question, or topic area, or phenomenon of interest. Individual studies contributing to a systematic review are called primary studies; a systematic review is a form a secondary study' (Kitchenham, 2004).

Both literature reviews and academic research papers have their similarities too. They contain some of the same elements. Usually, academic research papers contain a literature review section.

3. Method

The choice of methodology depends on the subject or field of the study. To establish the logical connection between the chosen methodology, the selected theory and the conclusions is a difficult task. The choice of methods may consist of three elements (Andersen, 1990):

1. Limit of empirical scope that should be studied;
2. The choice of a method for data collection;
3. Organizing of the exploration process.

A systematic literature survey has been conducted and this section describes the method used in the collection of the different publications to be surveyed. Also provided is the detailed information on how data was systemized and categorized.

3.1 A literature survey

This section gives an overview of the method of a literature survey as defined by different authors, and how a researcher should undertake it.

3.1.1 Recommendation by B. Kitchenham

A systematic literature review (SLR) consists of several activities. These activities can be grouped into three phases, as follows:

- Planning the review
- Conducting the review
- Reporting the review

Each phase has been deeply examined by Brereton et al. (2006) and Kitchenham et al. (2004). To get a full overview the reader can find it in Figure 4.

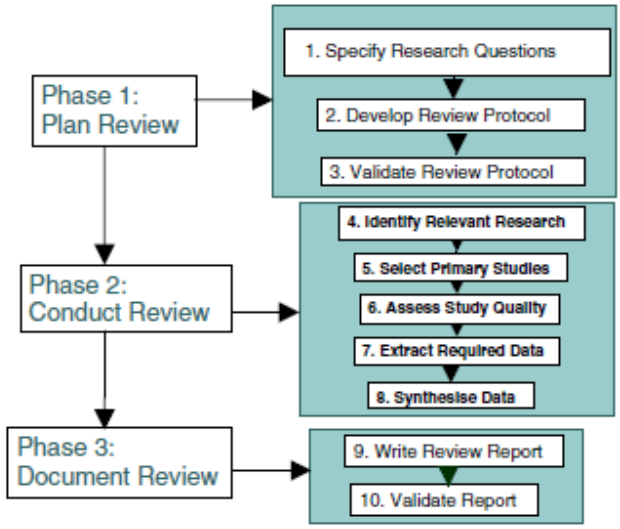


Figure 4: Systematic literature review process (Kitchenham et al., 2006).

According to Kitchenham et al. (2009) and Brereton et al. (2006), each phase can be broken down into detailed tasks, which are as listed:

- To review the existing guidelines for SLR across a range of domains.
- To identify the similarities and differences between them, in an attempt to propose appropriate guidelines to complete SLR.
- To provide a synthesis that is complete (with respect to defined criteria) and unbiased.
- To aggregate all existing evidence on a research question.
- To help the practitioners to use the guidelines, which provides human computer interaction in a specific context.

3.1.2 Recommendation by UNC

The Writing Center at University of North Carolina (UNC) at Chapel Hill (UNC Press, 2009) has an interesting and very useful website about literature review. This website was created to help both students and young researchers to get acquainted with the requirements for preparation and writing a SLR. Guidelines are written in a simple way, so that inexperienced researchers can pick up some ideas where and how they can start their own literature survey.

It is worth mentioning a couple of points from this website (UNC Press, 2009). Firstly, it states that a SLR is either a general description of articles published on a particular topic within a certain time period, or a simple summary of the sources. Generally, SLR has an organizational pattern, which combines synthesis and summary. A synthesis can be described as re-organization (or a reshuffling) of useful information, while a summary is just a recap of (going briefly through again) information that researchers have already described. The SLR evaluates the specific sources and at the same time gives advice to the reader or researchers on what information is the most pertinent.

Secondly, the website includes excellent guidelines about a paper's structure, dividing it into three parts, introduction, body and conclusion/recommendation. Each part is described in detail, and the authors give some examples. Because of this, it is easy to read and to learn from. The reader will receive advice on how to construct a working thesis statement, to use quotes sparingly, to revise the paper and to formulate citations.

3.1.3 Recommendation by A. Fink

According to Fink, a researcher might go through seven tasks to conduct a literature survey. These tasks are described as follows.

1. ***Selecting research questions*:** a researcher should have a research question before s/he starts work and this research question is a precisely stated question, which is meant to guide the review.

2. ***Selecting bibliographic or article databases, websites, and other sources:*** a bibliographic database is a collection of different articles, books, articles, etc. and this data helps the researcher to answer the research question. There are many different bibliographic databases, which can be classified by topic, such as biology, life sciences, medicine, business, engineering, computer science, etc. All these databases can be accessed online.

3. ***Choosing search terms:*** search terms are words and phrases and these guide the researcher in the search for relevant articles, books etc. It is very important that the researcher choose the right phrases. At the same time, it is important to use a particular grammar and logic (e.g. OR, AND) to conduct the search.

4. ***Applying practical screening criteria*:** literature searches usually result in a list of many articles, but on careful examination only a few articles might be relevant to the specific research question.

5. ***Applying methodological screening criteria*:** methodological criteria are criteria for evaluating the adequacy of a study's coverage.

6. ***Doing the review*:** reliable and valid reviews involve using a standardized form for abstracting data from articles, training reviewers (if more than one) to do the abstraction, monitoring the quality of the review, and pilot-testing the process.

7. ***Synthesizing the results:*** the results of the literature review may be synthesized descriptively and the review's findings are based on the reviewer's experience and the literature's content and quality.

Fink has written a description that is more detailed (Fink, 2009). An overview of all these steps may be found in Appendix F, Figure 23.

3.2 How articles were selected and analysed

There are many theories about the literature survey were presented above. Professor Fink (2009) presented in her book *Conducting Research Literature Reviews: From the Internet to Paper* that a researcher might go through seven tasks to complete a literature survey.

I decided to use this method written by Fink (see section 3.1.3) for this thesis, as it is straightforward, uncomplicated and clearly explained.

3.2.1 Selecting research questions

The main purpose of this thesis is to conduct a systematic literature survey of QoE research since 1999. I have focused on the following questions.

1) How many empirical studies were done in relation to 'Quality of Experience' in the ten years from January 1999 to June 2009?

2) How many empirical studies were undertaken, with human participants and without human participants?

3) Which aspects and sub-aspects of QoE have been evaluated in the empirical studies?

4) What types of study have been done by researchers and how do they relate to each other?

5) What are the findings in the chosen studies?

3.2.2 Search strategy: selecting databases

The search strategy consists of different electronic databases and the search terms that help to indentify the primary studies. These elements are discussed below.

I have identified seven electronic databases that I have used in my present thesis and these are listed in Table 1, which shows this in detail.

The column *Database* indicates the names of electronic databases.
The column *URL* shows the Uniform Resource Locator, i.e. the web address that each one has.

Database	URL
Google Scholar	http://scholar.google.com
ACM Digital Library	http://portal.acm.org
IEEE Xplore	http://ieeexplore.ieee.org
Science Direct	http://www.sciencedirect.com
Springer Link	http://www.springerlink.com
ISI Web of knowledge	http://apps.isiknowledge.com
HCI	http://hcibib.org

Table 1: Various relevant electronic databases with URLs.

3.2.3 Search strategy: choosing search terms

Initially, I tried several different search criteria, finally finding the acceptable criteria for use in this work (see Figure 18 in Appendix A).

I found that for this thesis it was better to use '**AND'** to conduct the search, because it gives the relevant hits from all seven databases.

To extract the primary studies, the search terms were used as following:

- **empirical study** and "**Quality of Experience"**[*]
- **1999** and **2009** (year)
- **Engineering, Computer Science, and Mathematics** (Subject Areas)

3.2.4 Applying practical screening criteria

It is of interest to note that the Google Scholar database gave the most hits, with 215 of 314 articles (68 per cent) but when I went through this list, I found that many studies were irrelevant to my chosen topic. In the case of all seven electronic databases (see Appendix C, Table 10), I had to remove some articles because some were duplicates.

One article I could not open, because of the broken link. Therefore, it was impossible to read the abstract and therefore the article was excluded from this thesis too. Another two or three articles were unavailable for free use. Since I could not read the abstracts and conclusions, I decided to exclude these articles as well. Some articles were also excluded from my survey because 'Quality of Experience' merely appeared in the articles' reference sections and QoE was discussed neither in the abstracts nor in the body of the texts.

Ultimately, this assessment resulted in 44 articles (see Figure 5).

[*] Note that I have use double-quote around Quality of Experience, because this combination gave me more relevant hits.

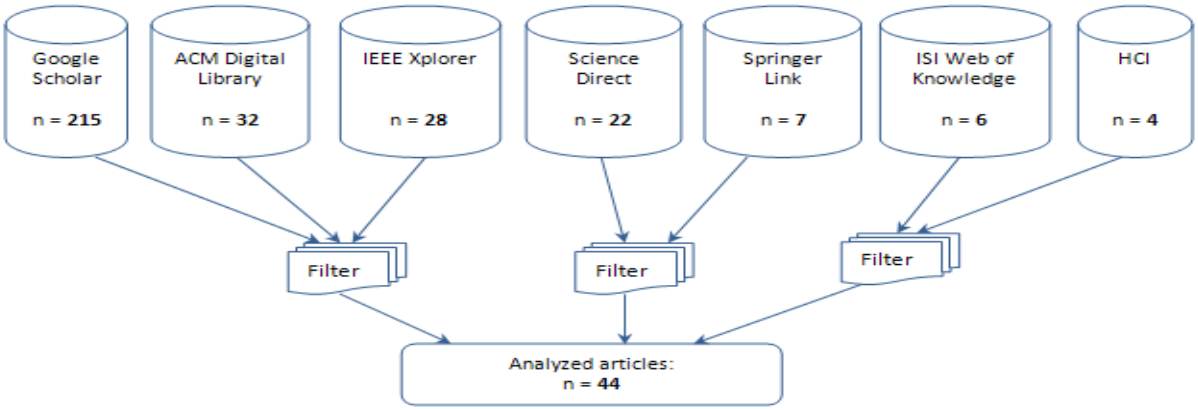


Figure 5: All databases show hits that were used in the present thesis.

A short overview of all analysed studies is given (see Table 16 and Table 17 in Appendix G).

3.2.5 Doing the review

First, I read both the title and abstract of each article. If it was not clear, I looked at the article's conclusion as well to find out more details. When I have got through all these articles I took 14 days' break. The reason for that was I wanted to go through all articles one more time, afresh, to make sure that all studies were correctly selected.

Furthermore, due to limited capacity and time, I had also decided that I was going to include neither master's theses nor PhD dissertations.

3.2.6 Synthesizing the results

All articles were carefully analysed by one researcher and details such as subject, type of study, aspects, purpose and results were written into an Excel spreadsheet. These details will be described in chapter 4 and chapter 5.

4. Analyses

The main point of this chapter is to give a thorough analysis of the selected empirical studies. In each study there are certain categories that are more interesting than others. Therefore, I decided to look at categories that are used or described in different studies from a Human-Computer Interaction (HCI) point of view. These categories are:

- *Subject*
- *Type of study*
- *Aspects*
- *Purpose*
- *Results*

Before I give the analysis of each category, it is worth making a note about the category '*aspects*'. I had thoroughly analysed some aspects reported by Nielsen, *A model of system acceptability* (Chapter 2.1.4), Alben *The ACM interaction design awards criteria* (Chapter 2.1.6), ISO 9126 standard (Chapter 2.1.2) and ISO 9241-11 Guidance on Usability (Chapter 2.1.3).

Aspects described in all these reliable sources do not fit my thesis one hundred per cent. Therefore, I have drawn a tree of my own aspects. Basically, I have used Alben's aspects model, which I have then expanded.

From the analysis of all the selected articles, I had picked up dozen of sub-aspects. All these sub-aspects have been assembled into small groups, which were then put into my version of Alben's aspects model. Figure 6 gives a clear picture of my framework and this is a short overview.

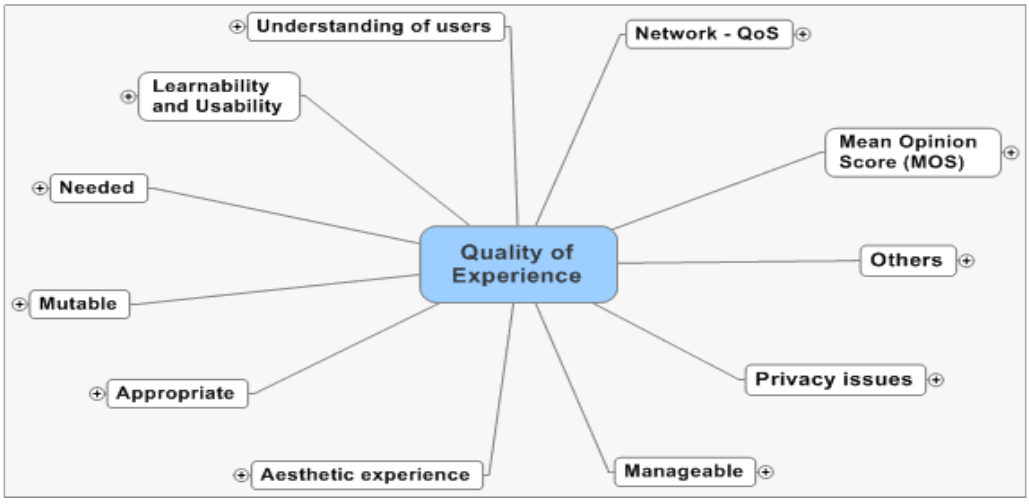


Figure 6 : A framework of 'Quality of Experience': a short overview.

Figure 6 shows that most of these aspect groups are identical with Alben's *The ACM interaction design awards criteria* model except for four aspects groups: 'network – QoS', 'Mean Opinion Score (MOS)', 'others' and 'privacy issues'.

Despite the fact that I was concentrating on researching the topic 'Quality of Experience', there are some empirical studies that have analysed both the phenomena 'Quality of Experience' and 'Quality of Service'. In these, the authors examined some factors/aspects such as jitter, delay, duration, packet loss etc. To fit these factors/aspects to my framework, I established the two groups '*network-QoS*' and '*Mean Opinion Score (MOS)*'.

There is one selected empirical study that has analysed sub-aspects such as 'music privacy' and 'private sharing'. Therefore, I decided to make a new aspects group, namely 'privacy issues'. There are also some other sub-aspects, which were not possible to place in existing aspect groups. Therefore, I had established the new aspect group 'others'. Finally, to obtain a full overview of my framework with all aspects and sub-aspects, the reader may look at Appendix F, Figure 24.

In total, 44 empirical studies were examined and a chronological overview of these studies may be found in Table 2.

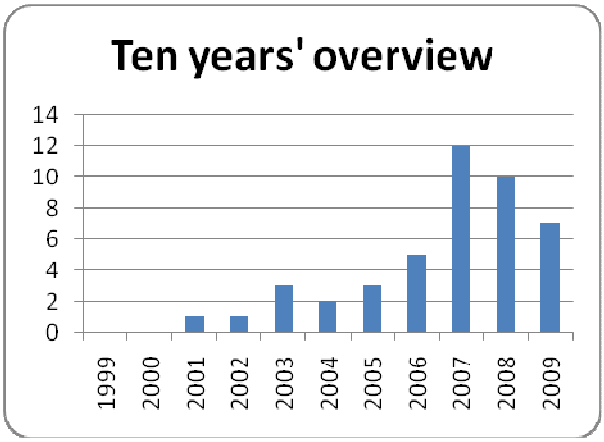


Table 2: Chronological overview of how many studies were published over a ten year period.

The first study was undertaken in 2001 and these data show that the phenomenon Quality of Experience is relatively new. Big interest in this topic started in 2007. The table shows that 12 studies were undertaken, which is almost thirty per cent of the studies undertaken during the ten year period. Only seven studies were undertaken in 2009. There is reason to believe that in that year more empirical studies might have been undertaken on this topic, because I finished the literature survey in the middle of July 2009.

4.1 Subjects

This section describes the subjects that took part in the studies investigated in the literature survey, the kind of information the researchers were able to obtain about them and how the participants were recruited.

The studies were split into two main groups: studies with human participants (e.g. school pupils, students, professionals, etc.) and studies without human participants (e.g. some studies undertaken with different types of topics such as wireless network, multimedia streaming, network traffic etc.).

The results show that 29 of the 44 studies (66 per cent) were conducted with human participants and 15 of the 44 studies (34 per cent) were conducted without human participants. A detailed analysis of these two main groups of subjects may be found in chapter 4.1.2, Table 4 and Table12 in Appendix E.

4.1.1 Studies with human participants

Table 3 presents:

1. The column **category of subjects** shows different kind of participants (e.g. students, professionals, etc.) in their experiments.
2. The column **experiments** is split into two columns, namely **N** and **%**. The former shows the number of studies and the latter shows the per centage of studies.
3. The column **participants** is split into three columns, namely **min**, **max** and **sum**.
   Min shows the minimum number of participants.
   Max shows the maximum number of participants.
   Sum shows the total sum of participants in different experiments.

Category of subjects		Experiments N	Experiments %	Participants Min	Participants Max	Participants Sum
TV channels	TV watchers (*)	1	3.4		297,600	297,600
Secondary school	School pupils	1	3.4	(20† -	120)	140
Students only	Undergraduates	5	17.3	20	515	829
	Graduates	2	6.9	20	62	144
	Subj. without numb.	1	3.4	x‡	x	x
	Type unknown	1	3.4	64	242	306
	Students total:	**9**	**31**	**20**	**515**	**1,279**
Professionals / Experts	a)	1	3.4	15	15	15
Mixed group of subjects	a)	3	10.4	6	148	175
Game players	a)	2	6.9	2	6	8
Others	a)	12	41.5	9	1,560	1,560
Total		**29**	**100**	**2**	**297,600**	**300,777**

Table 3: Overview of human participants in different studies.

Please note the letter a) indicates that more details about each category are provided in Table 12 (Appendix E).

In total, 300,777 subjects, who participated in 29 of the 44 empirical studies (66 per cent), were investigated in the literature survey. 9 of the 29 (31 per cent) were of the category of subjects 'students only', while 1of 29 (3.4 per cent) were of the category 'professionals /experts' and 12 of 29 (41.5 per cent) were of the 'others' category. More specific details of subject types divided into the categories can be found in Appendix E, Table 12. The number of participants per study ranges from 2 to 297,600 (see Table 3). There was one study with students with a lack of information about the amount of students who participated in the study.

The biggest category was 'TV channels', with 297,600 TV watchers, and it was 1 of the 29 studies involving human participants (3.4 per cent). This study was undertaken by using the most often watched 23 TV channels to find out more about two digital television technologies, namely IPTV and P2PTV, with a focus on content blocking probability. There is a lack of information about what kind of subjects participated in that study, therefore they can described only as 'TV watchers'.

* The case study was performed by using the most often watched 23 TV channels and there were 297,000 TV watchers who participated in this case.

† There is one empirical study that has been performed through two different tests. Twenty participants performed the first test and one hundred and twenty participants performed the second test. Therefore, I wrote in Table 3 that there are between 20 and 120 participants that participated in one study.

‡ The letter x means there is a lack of information about how many participants were involved in the study.

The category ‘students only’ was broken down into the different sub-categories ‘undergraduates’, ‘graduates’, ‘subjects without numbers’ and ‘type unknown’. The sub-category studied most often by researchers was ‘undergraduates’, who took part in 5 of the 29 studies (17.3 per cent) undertaken with human participants, while graduate students took part in 2 of the 29 (6.9 per cent). The category ‘students only’ has a total of 1,279 student participants.

Undergraduate students took part in 5 of 9 ‘students only’ studies (56 per cent of the student group), while ‘graduate’ students, ‘subjects without number’ and ‘type unknown’ all together took part in 4 of these 9 studies (the remaining 44 per cent).

The group ‘game players’ took part in two studies only and this group was the smallest of all. 2 game players participated in the first study and 6 participated in the second study. The total amount of participants of these studies was 8 and of the total it was 6.9 per cent.

Finally, the group delineated as ‘others’ included 1,560 participants. They took part in twelve experiments, and in total it was 41.5 per cent. For more information about this group, see Appendix E, Table12.

4.1.2 Studies without human participants

Studies without human participants are shown in Table 4, which consists of the following series of columns.

1. The column **topic** shows what kinds of topic authors have choosen in their studies.
2. The column **technology** shows what kinds of techniques or protocols authors or researchers have used in their studies.
3. The column **N** shows how many studies have been done.
4. The column **%** (per cent) shows what percentage this is of the topic.

Topic	Technology	N	%
Multimedia streaming	Audio and video streaming: H.264/AVC video format; video sequences such as Akiyo, Container and Foreman; voice-over-IP (VoIP); MPEG2 and MPEG-4 Part 10 format;	6	39,6
Network traffic	Gnutella peer-to-peer (P2P) network, Ethernet, UDP protocol, HTP techniques, TCP protocol, ADSL traffic traces,	4	26,8
Wireless network	IEEE 802.11*, IEEE 802.16, UMTS, WiMAX standards such as IEEE 802.16, IEEE 802.16a and IEEE 802.16e; WiFi, WLAN, GPRS, IEEE 802.11e HCCA; Multipoint to point IEEE WLAN 802.11a/g and Linksys WAP54G Access-Point for IEEE WLAN 802.11g	3	20,1
Web services	Three clustering Web services (J48 Classifier) and three clustering Web Services (Cobweb cluster).	1	6,75
Web search engine	User centric crawling to refresh schemes such as Staleness Based Refreshing (SBR) and Embarrassment-based Refreshing (EBR) methods.	1	6,75
Total:		**15**	**100**

Table 4. Overview of different topics and technologies in studies without human participants.

Table 4 shows studies that were conducted on five topics of interest to researchers: multimedia streaming, network traffic, wireless network, web services, and web search engine.

The most common topic was 'multimedia streaming', studied in 6 of the 15 studies (39.6 per cent). The topic 'network traffic' was studied in 4 of the 15 studies (26.8 per cent), while the topic 'wireless network' was covered by 3 (20.1 per cent) studies. The least common topics were 'web services' and 'web search engine', each covered by only 1 of the 15 studies (6.75 per cent).

Table 5 is an overview of the topics of studies without human participants, giving more information about each of the topics of Table 4.

1. The column **topic** shows what kinds of topics authors have chosen for their studies.
2. The column **description** describes the technologies that were used in different experiments. These technologies are for example video streaming using H.264/AVC video format, Voice-over-IP, MPEG2, wireless technologies such as IEEE 802.11, IEEE 802.16e, WiMAX, etc.
3. The column **N** shows the number of studies.

* Different wireless technologies, such as IEEE 802.11a, IEEE 802.11g, IEEE 802.16a, Worldwide Interoperability for Microwave Access (WiMAX), were used in various studies. For those who want to obtain more details, please look at Appendix D.

4. The column **observation** shows how many observations or objects were used in each study (e.g. nine different video sequences or twenty mobile nodes).
5. The column **%** (per cent) shows what percentage each topic, and each technology used, is of the total.

Topic	Description	N	Observation	%
Multimedia stream	Video and audio stream	2(*)	9	13,2
	Multimedia stream over IP	2	242	13,2
	Multimedia stream over Wireless network	2	18	13,2
	Total	**6**	**269**	**39,6**
Web services	Six services	1	6	6,75
Network traffic	Clients connected to DSLAM (†)	1	4,000	6,7
	Protocol (TCP, UDP and HPT)	1	9	6,7
	Gnutella peer-to-peer (P2P) network	1	92	6,7
	Simulate different traffic	1	3	6,7
	Total	**4**	**4,104**	**26,8**
Wireless network	Mobile nodes	1	40	6,7
	Different wireless network	1	3	6,7
	Wireless technologies used for IPTV	1	3	6,7
	Total	**3**	**46**	**20,1**
Web search engine	Boston Data set and Multiple site Data Set	1	35,000	6,75
Total		**15**	**39,425**	**100**

Table 5: Overview of topics of studies without human participants.

The total amount of observations was 39,425 in 15 of the 44 empirical studies (34 per cent).

'Multimedia stream' is split into three groups and the total amount of the observations in these studies were 269 (39,6 per cent).

The topic 'network traffic' was split into four small groups: 'clients connected to DSLAM', 'protocol (i.e. TCP, UDP and HPT)', 'Gnutella peer-to-peer (P2P) network' and 'simulate different traffic'. In total there were 4,104 observations in studies without human participants. The highest amount in this category belongs to the 'clients connected to DSLAM' group with 4,000, and each client was connected to only one DSLAM (Digital Subscriber Line Access Multiplexers). The group 'simulate different traffic' has fewer observations and Table 5 shows 3.

Researchers have undertaken three studies in the topic 'wireless network', which was split into three different categories. There are 46 observations. Two studies were used to test different

* One study does not contain information about the number of video and audio sequences.
† Digital Subscriber Line Access Multiplexers (DSLAM) before forwarding it through the local routers to the France Telecom IP backbone.

wireless network technology, while only one study with 40 observations belongs to the group ‘mobile nodes’.

Finally, the topic ‘web search engine’ has only one study and the total amount of observations was 35,000. The authors used two different data sets. The first one was the ‘Boston data set’, which consisted of approximately 16,000 web pages and the second one was the ‘multiple site data set’, which consisted of 19,000 pages.

4.2 Type of study

This chapter describes the type of study that takes part in related work and presents how researchers undertook their studies.

Studies with human participants

Figure 7 presents and describes different type of studies. 14 out of the 29 studies with human participants (48.2 per cent) were of the type ‘experimental study’, the most common type. The next most common was ‘survey’, being 9 of the 29 (31 per cent), while the least common type was ‘case study’, 6 out of 29 (20.8 per cent).

The type of study ‘survey’ can be broken down into two distinct sub-types: *survey* and *online survey*. 5 of the 9 surveys (55 per cent) are online surveys, where participants have had to complete one or more different tasks using the internet.

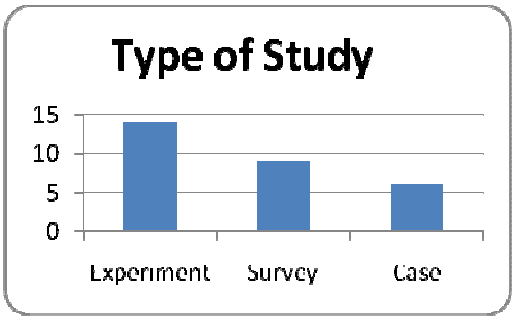


Figure 7: An overview of different type of studies with human participants.

Studies without human participants

Figure 8 presents and describes two different types of studies without human participants. A total of 14 out of 15 studies (93 per cent) were of the type ‘experimental study’. Only 1 of the 15 studies (7 per cent) was of the type ‘case study’.

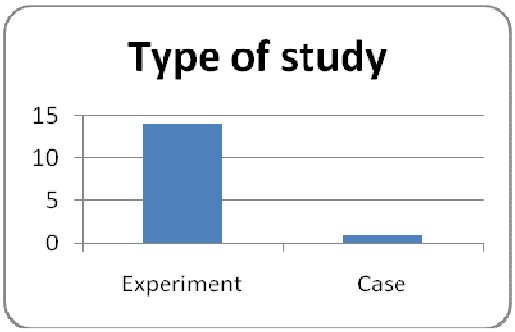


Figure 8: An overview of different type of studies without human participants.

4.3 Aspects

Table 6 presents all aspects from all the selected studies.

1. The column **author(s) and year** provides the surname of the first author and the year of publication.
2. The column **aspects** is split into 11 columns, each representing a different aspect group.

Please note that this table is only a short summary of the original table. It consists of four studies, while the full table is provided in Appendix F, Table 13.

Author(s) and Year	ASPECTS										
	Understanding of users	Learnability and Usability	Needed	Mutable	Appropriate	Aesthetic experience	Manageable	Privacy issues	MOS	Network - QoS	Others
Cheung et al., 2005		x*		x			x				
Mahlke, 2002		x				x					
Muntean et al., 2006		x									
Knoche et al., 2007							x				

Table 6: All aspects from all the selected studies.

* The letter x presents these aspect groups, which were described by the authors in each empirical study.

Using this table, it is possible to find out:

1. Which aspects were evaluated by a specific study.
2. How many different aspects groups were evaluated in each empirical study.

Figure 9 presents all the different aspects groups that were analysed in this thesis. The aspect group 'learnability and usability' was the most evaluated among scientists and in total 19 out of the 44 studies (43 per cent) covered it. The aspect group 'network – QoS' was evaluated in 15 out of 44 studies (34 per cent), while the aspect group 'manageable' was analysed in 14 of 44 (32 per cent) studies. Finally, there are a few other aspects groups that were analysed more than four times in the selected studies, namely 'aesthetic experience', 'mean opinion score', 'needed', 'mutable' and 'appropriate'.

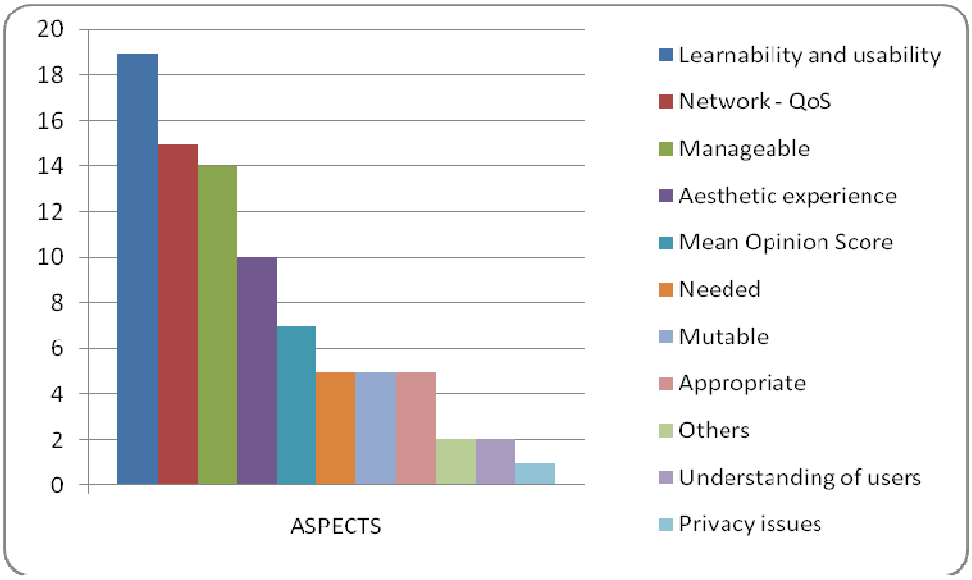


Figure 9: All the different aspect groups that were used in this literature survey.

It is interesting to note how many different aspect groups were used per individual study. There were five different aspects groups used in one study by Hamam et al. (2008), namely 'learnability and usability', 'appropriate', 'aesthetic experience', 'manageable' and 'network - QoS' (Table 13, Appendix F).

There are four studies that each used four aspect groups. These were Colbert, 2004; Zhang et al., 2007; Vegiris et al., 2008; and Navarro-Ortiz et al., 2008.

The most evaluated aspect groups were 'learnability and usability', 'network – QoS', 'manageable', 'aesthetic experience', 'mean opinion score', 'needed', 'mutable' and 'appropriate' (Figure 9).

Learnability and usability

Researchers were most interested in the aspect group 'learnability and usability' and this aspect group has thirteen sub-aspects. Figure 10 gives a clear picture of the sub-aspects.

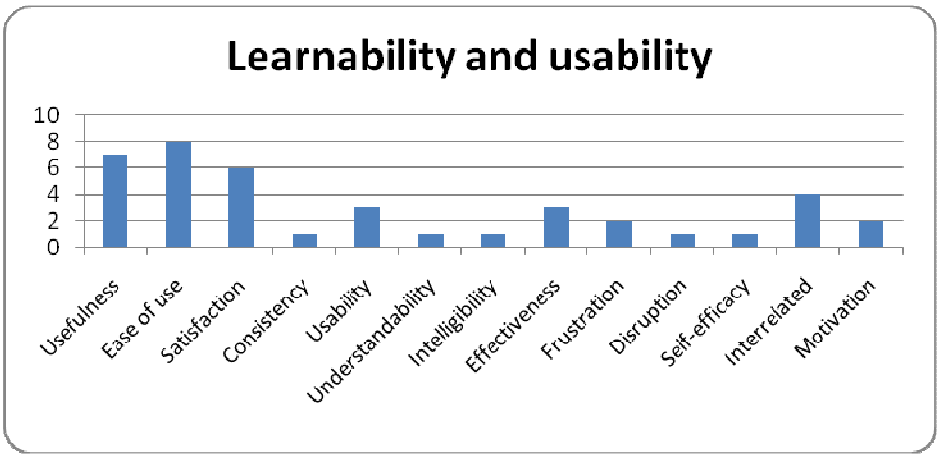


Figure 10: All the sub-aspects that belong to the aspect group 'learnability and usability'.

Please note that the sub-aspect 'interrelated' can be split again into the sub-aspects 'interactive', 'facilitation of role-playing', 'adaptivity' and 'utilitarian'. For a more detailed picture the reader should see Appendix F, Figure 24.

Figure 10 shows that the most evaluated sub-aspects in the studies with human participants are 'usefulness', 'ease of use' and 'satisfaction'. The sub-aspect 'ease of use' was evaluated in 8 of the 29 studies with human participants (27,6 per cent), while the sub-aspect 'usefulness' was evaluated in 7 of 29 (24 per cent) and 'satisfaction' in 6 of 29 (20,7 per cent).

Network – QoS

The aspect group 'network – QoS' was the next most evaluated among researchers and it was analysed in fifteen studies (see Figure 9). Network – QoS has thirteen sub-aspects such as 'ping', 'jitter', 'loss/packet loss', 'latency', 'delay', 'duration', etc. and it is show below in Figure 11.

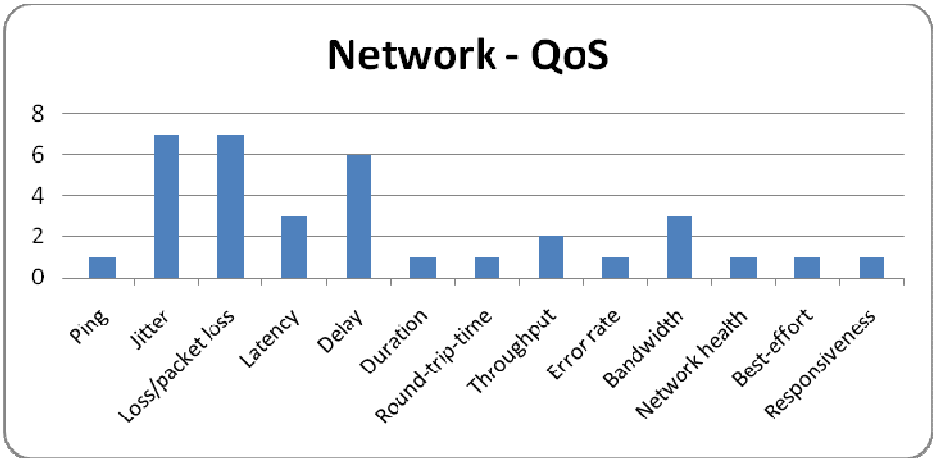


Figure 11: All the sub-aspects that belong to the aspect group 'network - QoS'.

Figure 11 shows the aspect group 'network – QoS' was analysed in studies without human participants. It shows the variation of each sub-aspect and the most analysed sub-aspects were 'jitter', 'loss/packet loss', 'delay', 'latency', and 'bandwidth'.

The most evaluated among scientist were the sub-aspects 'jitter' and 'loss/packet loss', each covered by 7 out of the 15 (46,6 per cent). The sub-aspect 'delay' was analysed in 6 of the 15 (40 per cent). Finally, the sub-aspects 'latency' and 'bandwidth' was described and analysed in 3 of the 15 (20 per cent) studies each.

Manageable

Authors in fourteen studies were interested in analysing the aspect group 'manageable' (Figure9). Figure 12 gives the details of the eleven sub-aspects. The sub-aspect 'practical acceptability' is split again into two sub-aspects, 'reliability' and 'acceptability'. More details about each aspect can be found in Appendix F, Figure 24.

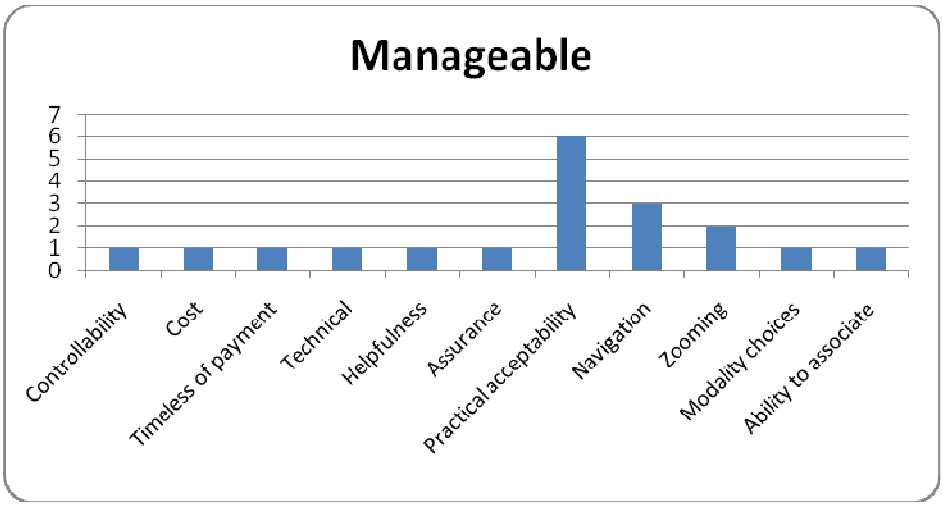


Figure 12: All the sub-aspects that belong to the aspect group 'manageable'.

Figure 12 reveals which was the most evaluated sub-aspect in the diverse studies. The most studied was 'practical acceptability', which was evaluated in 6 of the 29 studies that had human participants (20,7 per cent). As mentioned above, the sub-aspect was split again into two sub-aspects, 'reliability', which was evaluated in four studies, and 'acceptability', which was evaluated in two. The sub-aspect 'navigation' was analysed in 3 of the 29 (10 per cent). Finally, the sub-aspect 'zooming' was described and analysed in 2 of the 29 (2,9 per cent).

Aesthetic experience

The aspect group 'aesthetic experience' was often evaluated (Figure 9), and there were ten studies that did so. Figure 13 presents its thirteen sub-aspects. One of the sub-aspects, 'emotions', has sub-aspects of its own, namely 'conviviality (positive)', 'trust', 'irony', 'lies', 'conversational', etc. More details about this are provided in Figure 24, Appendix F.

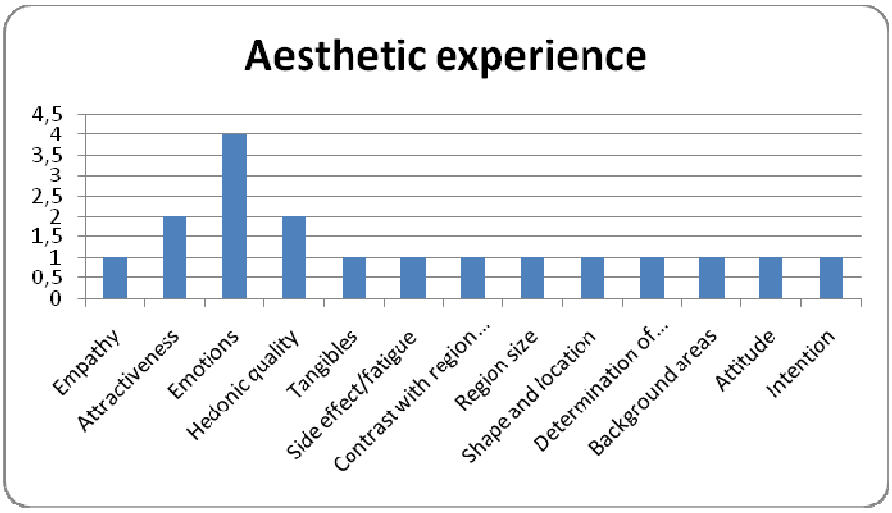


Figure 13: All the sub-aspects that belong to the aspect group 'aesthetic experience'.

Figure 13 shows that the most evaluated sub-aspects are 'emotions', 'attractiveness' and 'hedonic quality'. The sub-aspect 'emotion' was used in 4 of the 29 studies (13,8 per cent), while the sub-aspects 'attractiveness' and 'hedonic quality' were studied in 2 of the 29 (7 per cent).

Mean Opinion Score

The aspect 'mean opinion score' is less common than others (Figure 9); it was used in seven studies. Figure 14 presents its fourteen sub-aspects. It is worth noting that the sub-aspect 'zapping (effect)' was split again into the four sub-aspects 'round-trip delay', 'zapping time', 'packet losses' and 'video quality'. More details about this is provided in Figure 24, Appendix F.

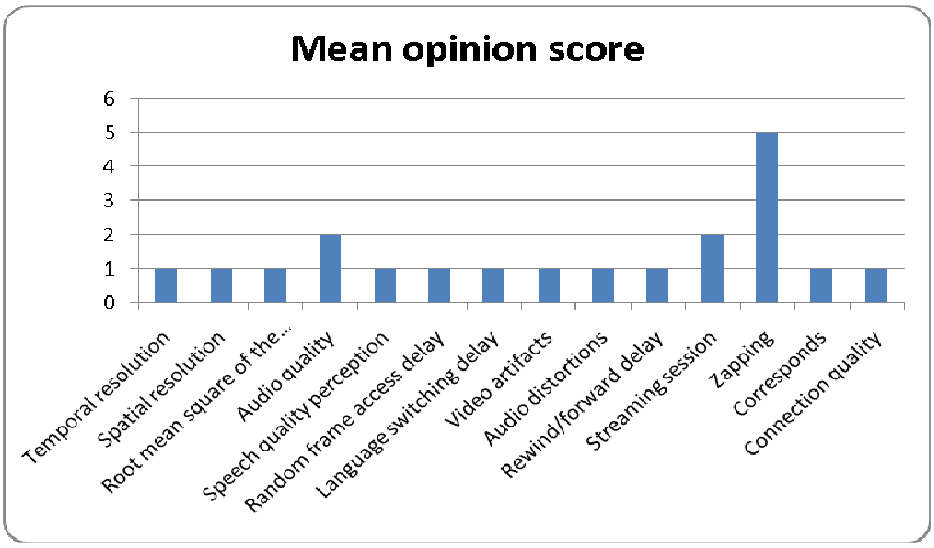


Figure 14: All the sub-aspects that belong to the aspect group 'mean opinion score'.

Figure 14 shows each sub-aspect. Note that the sub-aspect 'zapping' was used in five different studies, while the sub-aspects 'audio quality' and 'streaming session' were evaluated in two.

Needed

The aspect 'needed' was of less interest to researchers (Figure 9); it was in just five studies. Figure 15 presents its five sub-aspects, as used in studies by various researchers.

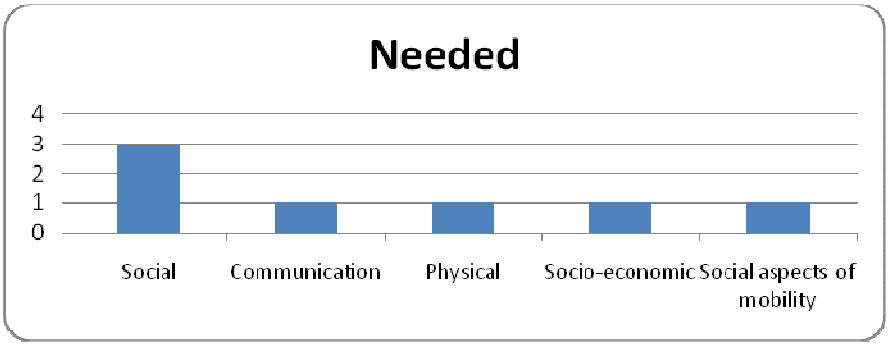


Figure 15: All the sub-aspects that belong to the aspect group 'needed'.

Figure 15 shows that the sub-aspect 'social' was evaluated in three studies, while the other sub-aspects were each evaluated in only one study.

There are also less-studied aspect groups such as 'others', 'understanding of users' and 'privacy issues'. These aspect groups were of low interest and less than five studies have analysed each aspect group. Therefore, I do not need to present them here and refer the reader to Appendix F, Figure 24 for more information.

4.4 Purpose

Table 7 presents an overview of the purpose of all of the studies.

1. The column **author(s) and year** provides the surname of the first author and the year of publication.
2. The column **purpose** describes the purpose of each study.

Author(s) and Year	Purpose
Cheung et al., 2005	To examine the assymetrical effects of negative and positive website attribute performance on satisfaction.
Mahlke, 2002	Examines the role of subjectively perceived factors of the experience of website usage in forming an intention to use a website.
Muntean et al., 2006	The aim of the experiment presented here was to investigate the usability and effectiveness of novel QoE layer enhancement for Adaptive Hypermedia Systems (AHS), in the educational area.
Knoche et al., 2007	To determine the optimal zoom factors depending on the target device size in terms of its native resolutions. To give the participants a choice of following two clips playing in parallel by watching either the left or the right half of a screen.
Colbert, 2004	To compare students' experience of communication before and during a rendezvous.

Table 7: An overview of the purpose of all of the studies.

The reader might notice that this table is a short summary of the original one. Here, only five empirical studies are listed and the aims of the studies are provided, but the complete table can be found in Appendix F, Table14.

The table presents raw data, but for the reader who is not familiar with the all studies, it is not easy to interpret the data. Therefore, I decided to create some categories. These categories are: 'education and communication', 'multimedia stream', 'real-time collaboration', 'resource sharing and trustworthiness', 'size and resolution', etc. (see Figure 16 below). A short synopsis of these categories is given below.

Education and communication

Ten studies in this category were undertaken on education and communication (Colbert, 2004; Muntean et al., 2006; Ras et al., 2007; Collange et al., 2008; Agrawal et al., 2007; Muntean, 2007; Schaik et al., 2003; Koivisto, 2007; Velsen et al., 2006; Siller et al., 2006;). The first study compared students' experience of communication before and during a rendezvous (Colbert, 2004). The authors rated several aspects of the experience of communication and students rated them lower during a rendezvous than before a rendezvous. The second study tested two groups of students (Muntean et al., 2006). Using two different systems, the students' information processing time per page was tested, and how quickly they learn from websites was tested. The other eight studies have to do with either learning or communication.

Multimedia streaming

This category contains nine studies, which have analysed multimedia streaming, i.e. audio or video (Siller et al., 2003; Waltermann et al, 2008; Muntean et al., 2007; Lu et al., 2008; Monteiro et al., 2007; Lopez et al., 2006; Ciubotaru et al., 2009; Calyam et al., 2007; Moid et al., 2009;). One study examined the performance of multimedia streaming over a wireless network, i.e IEEE 802.11b Wireless WLAN, while another study used four short video sequences in an experiment. The students analysed these sequences giving attention to problems dealing with visual deficiencies.

Real-time collaboration

This category consists of nine studies, which analysed users' experiences in real time or real video transmissions (Wattimena et al., 2006; Silzle, 2007; Holone et al., 2008; Vegiris et al., 2008; Dyck et al., 2004; Calyam et al., 2008; Zhang et al., 2009; Navarro-Ortiz et al., 2008; Yoo et al., 2008;). One study undertook four real-time transmissions from the Thessaloniki Concert Hall (Vegiris et al., 2008). This experiment was set up in four cities in Greece (Athens, Thessaloniki, Patras and Heraklion) and one elsewhere in the EU (Dublin, Ireland). The project involved seven virtual halls, which were different in size, from very small to medium-sized. The main focus of this study was to give theatre audiences the same user experience, despite the fact that they were sitting in virtual halls.

Resource sharing and trustworthiness

This category contains five studies, which analysed technologies such as Gnutella peer-to-peer file sharing network or the online distribution of digitized goods (Bhattacharjee et al., 2003; Ali et al., 2005; Liu et al., 2006; Dufour et al., 2006; Hamam et al., 2008;). One study undertook experimentation on revenue and piracy implications of legal online music. Its authors wrote, 'The music industry is one such business sector that has been dramatically affected by the online sharing phenomenon' (Bhattacharjee et al., 2008). Dufour et al. (2006) analysed the Gnutella peer-to-peer (P2P) file sharing network performance of queries and queries hits by using neighbor selection algorithm. The main focus of this study was to show that this algorithm improve P2P performance and provide better QoE.

Size and resolution:

This category contains two studies, which analyse optimal zoom depending on the product display size (Knoche et al., 2007; Knoche et al., 2009;). Mobile TV has a small display size, therefore zooming plays an important role. In the future, designers have to know what size of screen dimensions will achieve user satisfaction when devising screen sizes for devices like Mobile TV and the iPhone, etc.

Figure 16 gives an overview of the purpose of all the analysed studies in the present thesis.

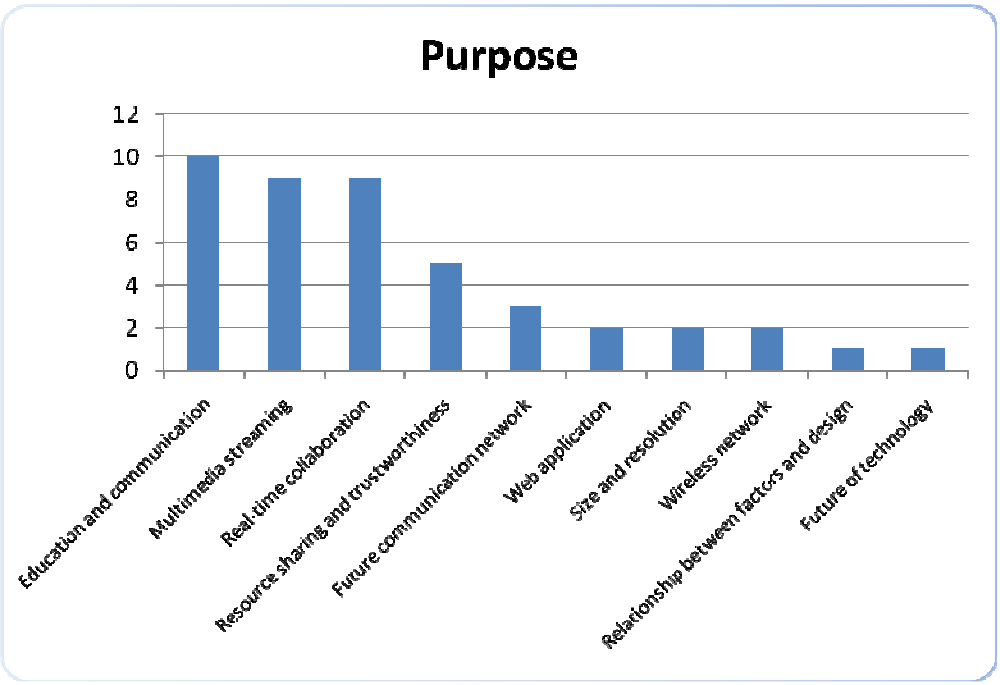


Figure 16: An overview of the purpose of all studies.

The results show that in 10 of the 44 studies (23 per cent) the most frequently evaluated category was 'education and communication'. From these results, it is obvious that education and communication were important to the researchers.

The category 'multimedia streaming' was analysed in 9 of 44 studies (20 per cent). The interest in this category was both audio and video streaming. The scientists were interested in the H.264 video format, which has been recently adapted in mobile broadcasting and other video streaming networks. They analysed how to achieve good balance between video quality and packet processing time. Other studies have used a few video clips to analyse them with participants.

The category 'real-time collaboration' was analysed in 9 of the 44 (20 per cent) studies, using new technologies i.e. remote access of electron microscope, virtual chat-room application, live theather via internet etc. For example, the application 'E-drama: online role play' is a virtual environment where people can communicate with each other online, for example, a 3D graphic environment giving participants an experience of a 'real' world (Zhang et al., 2009).

4.5 Results

This section presents the results of this literature survey to show what kind of information the authors were interested in and what the findings of their studies were.

All results, which have been systematically and carefully analysed for each article separately, are included in Table 8 (below). I have used two main columns in this table.

The column **author(s), year** provides the surname of the first author and the year of publication. The column **results** provides a short synopsis of the results of all studies.

Author(s) , Year	Results
Cheung et al., 2005	Provide partial support for the asymmetric effect on online user satisfaction. The positive performance of understandability, usefulness, and navigation had greater power than their negative performance, whilst the negative performance of reliability, access, and usability had a greater impact than their positive performance.
Mahlke, S. 2002	The four aspects of experience (i.e. usefulness (U), ease of use (EOU), hedonic quality (HQ) and visual attractiveness (VA)) succeed all in predicting the intention to use. The weight of contribution to the intention to use differs from U with the major influence to VA with the smallest effect.
Muntean et al., 2006	• Students from the QoE Adaptive Hypermedia Systems (QoEAHS) group had shorter *Study Session Time* than those that used the AHA! system. Therefore, the QoE-aware system has ensured a smooth learning process • QoEAHA system provided better end-user satisfaction, than the AHA! system. It shows that the students considered the QoEAHA significantly more usable than AHA! • There is no strong correlation between learning performance and perception of usability • An overall usability assessment shows that students considered QoEAHA significantly more usable than AHA!.

Table 8: A short synopsis of the results from each article.

The reader might notice that this table is a short summary of the original. Here just three empirical studies are presented, together with the authors' results, but the complete table can be found in Table15, Appendix F.

I have found many different aspects and sub-aspects that were discussed by the authors, therefore I decided to highlight those that were discussed in more than three articles. Figure 17 shows which sub-aspects the authors were most interested in.

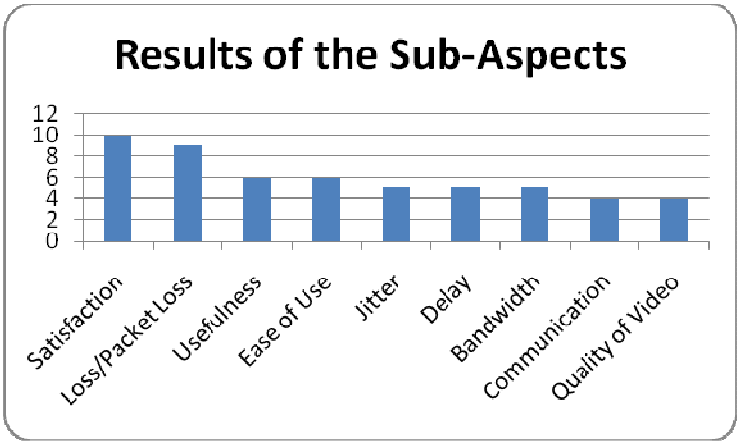


Figure 17: Results of the sub-aspects, picked out from the analysed articles.

Figure 17 shows that the sub-aspect that was studied most often by researchers was 'satisfaction'. Ten studies analysed this sub-aspect. The next most studied sub-aspect was 'loss/packet loss', covered by nine studies. The sub-aspects 'usefulness' and 'ease of use' were each analysed in six studies, while 'jitter', 'delay' and 'bandwidth' were each in five. Finally, the sub-aspects 'communication' and 'quality of video' were each analysed in four studies.

Sub-aspect: **Satisfaction**

'Satisfaction' as shown in Figure 17 was the most often analysed sub-aspect. Cosma et al. (2008) have undertaken a study of the new technology of wireless IEEE 802.16 (WiMAX) client performed on audio and video streams. The results were satisfactory for audio streams, but for video streams there were some interruptions and image 'blockiness'. The study shows that the IEEE 802.16 client lacks reliability.

There are three studies about learning achievement: Muntean et al., 2006; Muntean, 2007; and Hamam et al., 2008. The findings from the first two studies are that those students who belong to the QoE Adaptive Hypermedia Systems (QoEAHA) group performed significantly better in learning than the AHA! group. This means that the QoEAHA system provided a better QoE for the learners and that students from this group achieved satisfaction. Hamam et al. (2008) have undertaken a study with the Haptic User Interface (HUI) application and the results show that it is possible for researchers to measure the quality of experience of the users.

The statistical method Cronbach's Alpha was used to measure different sub-aspects. The results show that 'satisfaction' has a value of 0.84 and this indicates an acceptable level. This study was undertaken by Zhang et al. (2007).

Finally, the sub-aspect 'satisfaction' was of interest to Velsel et al.(2007). The results from their study showed that organizations should pay attention and provide a high quality of service, with a helpline or helpdesk to achieve satisfied and motivated customers.

Sub-aspect: **Loss/packet loss**

The sub-aspect 'loss/packet loss' has been investigated in several studies (Wattimena et al., 2006; Waltermann et al., 2008; Muntean et al., 2007; Collange et al., 2008; Lopez et al., 2006; Dyck et al., 2004; Garcia et al., 2009; Navarro-Ortiz et al., 2008; Cosma et al., 2008).

Cosma et al. (2008) found the wireless client Wifi 802.11 performance to be very good and it had 0% packet loss for audio stream, while the video stream had 0.034% packet loss. The results showed that if customers used this technology, the Quality of Experience (QoE) was excellent on both audio and video transmissions (Cosma et al., 2008).

Another study showed that different wireless technologies behaved differently. The client Wifi 802.11a was less robust that WiMAX technology, which proved to be very robust and flexible (Garcia et al., 2009).

Some researchers found that the sub-aspect 'latency' of UDP-based techniques was not affected significantly by the sub-aspect 'loss', but it depended on bandwidth (Dyck et al., 2004). An interesting finding from other researchers was that the factor 'response time' was significantly influenced by the retransmission delays of lost packets. The authors proved the correlations between some traffic features and the network performance, i.e. the higher the loss rate, the shorter and smaller the TCP connections (Collange et al., 2008).

Finally, other researchers were interested in the performance of adaptive multimedia streams through wireless technology. Muntean et al. (2007) has found that for the same number of simultaneous multimedia clients, the average values of the performance sub-aspects i.e. loss, delay, and total throughput were better for Quality-Oriented Adaptive Scheme (QOAS) than for other solutions i.e. Loss-Delay Adaption Algorithm (LDA+) and TCP-Friendly Rate Control Protocol (TFRC). Furthermore, the results indicate that for the same number of clients QOAS was possible to achieve a higher average end-user quality. The findings indicate that by using QOAS an increased level of QoE can be provided to each viewer.

Sub-aspect: **Usefulness**

All six studies (Figure 17) described 'usefulness' as being important (Cheung et al., 2005; Mahlke, 2002; Holone et al., 2008; Morris et al., 2001; Koivisto, 2007; Pianesi et al., 2009). Pianesi et al. (2009) in their study found that the driving force for acceptance was a guide's 'usefulness' in improving interaction with the particular museum. Furthermore, Koivisto (2007) wrote in his results that the sub-aspect 'usefulness' was of more importance than 'ease of use' in the acceptance of basic communication services. Morris et al. (2001) found that those two sub-aspects, 'ease of use and 'usefulness', are very significant determinants of IT use.

A few researchers have made investigations into 'usefulness' of service. Holone et al. (2008), in the study of a particular system, found that the trend of the 'usefulness' ratings was positive and that there were some users who rated the 'usefulness' of the system very highly. Furthermore, the authors wrote, 'Interestingly, even when participants rate the usefulness lower, they tend to think about it as a useful service for others'.

Mahlke (2002) found that 'usefulness' succeeds in predicting the intention to use a website. A questionnaire was developed to measure the 4 sub-aspects of experience and the intention to use a website. 79% of the total variances were explained by the sub-aspects 'usefulness' (U), 'ease of use' (EOU), 'hedonic quality' (HQ) and 'visual attractiveness' (VA). Furthermore, the author found that the intention to use a website was most affected by U, while VA had the least effect.

Sub-aspect: **Ease of use**

There are six studies (Figure 17) that have analysed 'ease of use' as an essential sub-aspect (Mahlke, 2002; Hamam et al., 2008; Koivisto, 2007; Schaik et al., 2003; Morris et al., 2001; Pianesi et al., 2009). Mahlke (2002) found that this sub-aspect contributes significantly to the intention to use a website. 'Ease of use' is a sub-aspect used for a better understanding and explanation of the concept of user experience. The experiment shows variations in the effect of aspects depending on the time and importance of various aspects, which in turn depend on the person variable, e.g. the level of internet expertise of an individual person.

Hamam et al. (2008) have used the haptic learning system and the haptic enabled UML CASE tool. Results show that this system is able to measure the QoE of the participants. The authors have also noticed also that Haptic-based GUI is quite a new application domain. Therefore, further research on this system needs to be undertaken, to achieve better results according to the QoE of a haptic user interface (HUI) application.

Koivisto (2007) undertook a study in which some aspect groups such as 'ease of use', 'usefulness', and 'attitude' were analysed. Three different groups of students, Finnish, Spanish and Chinese, were tested. The results showed that the sub-aspect 'ease of use' was less important than 'usefulness' in the acceptance of basic communication services in community context.

Schaik et al. (2003) used two methods, the Likert scale and the visual analogue scale, to measure different aspects. The results showed that the statistical method Cronbach's Alpha, when used to measure 'ease of use' value, was acceptable (0.89). This shows that Likert and the visual analogue scale formats are reliable and respectable. Pearson's correlation method was also used to establish validity. The correlation between 'intensity of flow', 'ease of use' and 'disorientation' was tested on the visual analogue scale format. The results showed that the correlations were moderate and significant, while there was a lack of correlation between 'ease of use' and 'disorientation'. The correlation was also tested between 'ease of use' and experience in using the Web and the results showed that 'ease of use' was significant on the visual analogue scale format. A point mentioned by the authors is that the participants themselves argued for the sub-aspect 'ease of use' as a particular benefit.

Sub-aspect: **Jitter**

The sub-aspects 'jitter' has been investigated in several studies (Cosma et al., 2008; Garcia et al, 2009; Dyck et al., 2004; Waltermann et al., 2008; Wattimena et al., 2006).

Cosma et al. (2008) analysed the statistical interpretations of three metrics, One-Way Delay (OWD), Packet Loss Ratio (PLR) and Inter-Packet-Delay Variation (IPDV). Their results showed that the value 'jitter' of the ethernet client delivered a high QoE for both audio and video transmissions.

Garcia et al. (2009) tested different wireless technologies, such as IEEE 802.11a, IEEE 802.11g and Worldwide Interoperability for Microwave Access (WiMAX). The findings were that the average 'jitter' was very similar in WiMAX and IEEE 802.11a. According to the authors, these technologies are approximately 35% better than IEEE 802.11g. One of the reasons for that is that the wireless technology IEEE 802.11g seems to be less stable. All measurements were made using an IEEE802.11b/g network with many Access Point (APs) that could interfere with the IEEE 802.11g deployment. IEEE 802.11g would probably obtain a better network's QoE parameter without this university network.

Dyck et al. (2004) undertook a study in which they analysed a telepointer, which is valuable for supporting real-time collaboration, but it is not often used in commercial groupware applications that run on the internet. The findings showed that when not predicting, error arises from jumps in the telepointer path due to lower send rates, 'jitter' and loss.

Waltermann et al. (2008) examined a new technique for seamless VoIP-Codec Switching in Next Generation Networks vertical handovers and this study shows that sub-aspects 'jitter' has negative effect on QoE. The results showed that in 90% of the test cases, no interruption occurred in the audio stream during the switching period. The quality gain of 5.4% was estimated for the scenario. The results showed that the procedure was constrained by the network-inherent packet loss, and the 'jitter' buffer status.

A study was undertaken examining internet *Quake* players. Wattimena et al. (2006) found that 'jitter' and 'ping' have a significant negative effect on QoE. This effect occurs on the subjective 'Mean Opinion Score (MOS)' and objective gaming quality. Furthermore, the authors emphasized that 'jitter' in the network had a very negative effect on the *Quake IV* players' QoE.

Sub-aspect: **Delay**

There are five studies (Figure 17) described 'delay' as being significant (Muntean et al., 2007; Lopez et al., 2006; Moid et al., 2009; Navarro-Ortiz et al., 2008; Cosma et.al., 2008).

Different metrics, such as the abovementioned One-Way Delay (OWD), Packet Loss Ratio (PLR) and Inter-Packet-Delay (IPDV), were analysed by Cosma et al. (2008). They found that audio and video transmissions delivered a high QoE with the low delay value of the ethernet client.

The solution for the IEEE 802.11e Hybrid Coordination function Controlled Channel Access (HCCA) mechanism, which has been designed for a heterogeneous network, was analysed by Navarro-Ortiz et al. (2008). The findings showed that VoIP maintains good quality (MOS is approx. 4.5). Furthermore, the authors analysed the Monolithic Shaper-Scheduler (MSS) with the territory method. When the territory method was used, the quality perceived by a conversational user was less influenced by the traffic generated by other services. End-to-end impairment, i.e. 'delay', was considered. The IEEE 802.11e wireless network reacted against the impairment, maintaining the demanded quality. Another issue was that the QoE, which is necessary for the different traffic classes, was found to be customizable.

Moid et al. (2009) tested the three video sequences Akiyo, Container and Foreman[*]. The results showed that adapting the transcoding parameters at the application layer to the current channel condition did not produce significant processing delays.

In one study, undertaken by Lopez et al. (2006), the behaviour of multimedia systems delivered over IP networks was characterized. One point from this study was that the most prominent sub-aspect affecting the observed response was 'delay' and this factor was statistically significant. Another issue is that ANOVA analysis revealed that 'delay' play an important role for the response variables such as QoEaudio and QoEvideo. According to the R-squared value, the model fits the behaviour of the response variable QoEvideo quite well. The results showed that the factor 'delay' affected the final video quality QoEvideo.

Sub-aspect: **Bandwidth**

There are five studies that have analysed the sub-aspect 'bandwidth' (Monteiro et al., 2007; Dyck et al., 2004; Calyam et al., 2008; Garcia et al., 2009; and Cosma et al., 2008;). One study described the impact of network connection quality on the user QoE MOS in a remote microscopy session. This study was undertaken by Calyam et al. (2008). The results show that the real-time control and video image transfer traffic depends on bandwidth intensity for achieving 'at-the microscope' QoE.

Garcia et al. (2009) analysed three wireless technologies: IEEE 802.11a, IEEE 802.11g and WiMAX. The finding was that effective bandwidth, i.e., the IPTV content distribution system, could support 5 IPTV channels in the WiMAX scenario. Furthermore, they studied QoE comparison between wireless technologies and the results showed that the IEEE 802.11 technology had a very stable network's QoE.

Sub-aspect: **Communication**

The sub-aspect 'communication' was analysed in four empirical studies (Colbert, 2004; Ras et al., 2007; Agrawal et al., 2007; Zhang et al., 2009; and Koivisto, 2007;). Zhang et al. (2009) illustrated a multi-user role-playing environment 'e-drama'. This system provides a possibility for developers to find out how emotional issues can be used in different scenarios and how people can find new ways to communicate with each other. Their results showed that new systems provide new opportunities, that is, new channels of communication between, for example, children or young people. This system also provides opportunities for professional training, where face to face meetings can be expensive or impossible, e.g. e-learning in the workplace. This system also shows that the system may be used in normal classroom education, providing 24/7 efficient personalized social skills, career training via role-play and good communication.

Ras et al. (2007) used a free and open-source Wiki-based system called software organization platform (SOP) to support and manage both students' experiences and observations. The results showed that communication is an essential aspect of SOP. Wiki functions also support students' information-sharing.

[*] There are three test video sequences, i.e. Akiyo, Container and Foreman, all of which were used in the experiment.

Colbert (2004) analysed users' experiences of communication before and during a rendezvous. The results showed that the users' experience was better before a rendezvous than during it. The communication during a rendezvous was worse in the sub-aspect 'social acceptability' (mean rate before = 4.16, sd = 0.97; during = 3.86, sd = 1.19; p = 0.001), the sub-aspect 'frustration' (mean rate before = 1.41, sd = 0.77; during = 1.62, sd = 1.05; p = 0.011) and the sub-aspect 'disruption' (mean rate before = 1.41, sd = 0.74; during = 1.68, sd = 1.02; p = 0.001). The results also showed that context-aware or mobile communication is important for overcoming the barriers of 'disruption' or 'frustration'. The 'satisfaction' sub-aspect is not necessarily reduced by additional so-called complication of mobile communication.

Sub-aspect: **Quality of video**

There are four studies that analysed the sub-aspect 'quality of video' (Agrawal et al., 2007; Ciubotaru et al., 2009; Vegiris et al., 2008; Knoche et al., 2009). Ciubotaru et al. (2009) described the Region of Interest-based Adaptive Scheme (ROIAS), which adjusts differently the regions within each frame of the streamed multimedia content, which is in turn based on the users' interest in them. Ciubotaru et al. (2009) presented two objective metrics, which estimate user-perceived video quality, such as Pick Signal-to-Noise Ration (PSNR) and Video Quality Measurement (VQM). These metrics show how by using ROIAS there is a clear benefit in terms of quality in the areas, which are of hight interest for the user.

Vegiris et al. (2008) undertook another study and the results showed that users evaluated the quality of video as 'very good' or 'excellent' and it was more than 80%. The users evaluated the quality of sound to be more than 70%.

It was pointed out that lab experiments may be a conservative estimate of the acceptability of video quality as consumed by people on the move, e.g. on a train (Knoche et al., 2009).

5. Discussions and Summaries

This chapter discusses the findings described above. Firstly, the results of studies with human participants (Chapter 5.1) will be presented. Then, the results of studies without human participants (Chapter 5.2) will be presented. At the end, a short summary will be given.

5.1 Studies with human participants

The sub-aspects 'satisfaction', 'usefulness', 'ease of use' and 'communication' belong to empirical studies that have been conducted with human participants (Figure 14., Chapter 4.5).

As described in chapter 4.1.1, Table 3 presents the category of subjects 'TV channels'. Lu et al. (2008) used the largest group (the largest amount) of participants, namely 297,600 TV watchers. This study compared two technologies, IPTV and P2PTV, to find out which of these technologies incurs the lowest end-to-end blocking.

The findings indicate that when the amount of users increase, there will be a point at which the blocking in P2PTV will be less than in IPTV, unless the IPTV network is extended accordingly. In our opinion, these findings are essential to P2PTV network. Nowadays, there are many households or organizations that are more or less related to these technologies, i.e. IPTV and P2PTV. QoE can be achieved when techniques and services work properly. Therefore, it is important to choose the right technology or service to satisfy user experience. This is only a single study undertaken in this field, so more research on these two technologies is needed in order to confirm or disprove the findings.

Furthermore, Table 3 presents another category of subjects, namely 'students only'. In this category there are five studies where the sub-aspect 'satisfaction' was analysed: Cheung et. al. (2005); Muntean et al. (2006); Colbert, (2004); Velsen et al. (2007); and Muntean et al. (2007). Two studies – Muntean et al. (2006) and Muntean (2007) – experimented on this sub-aspect and the results showed that by integrating learning systems, like in the QoE Adaptive Hypermedia System (QoEAHS), a better learning performance may be achieved. The studies also showed that the Adaptive Hypermedia System provided a better QoE for the learners, students, and participants for gaining 'satisfaction'. Another two studies – Colbert, (2004), Velsen et al. (2007) – experimented on the sub-aspect 'satisfaction' and showed that, through relationships between people, users can achieve satisfaction. The results showed that user experience does not appear to differ in terms of satisfaction and that there is a positive relation between the overall help facility quality and the customer's satisfaction. The last study – Cheung et al. (2005) – experimented on

the asymmetric effect of websites. This study is different from the other four studies mentioned above: the results showed that the sub-aspects 'reliability', 'accessibility', and 'usability of website' must function very well, otherwise users' satisfaction will drop dramatically. Furthermore, the authors suggested that the sub-aspects 'understandability', 'usefulness', and 'navigation' were motivating factors. When these motivating factors find a place in websites, then user satisfaction will jump significantly. According to these findings, we can conclude that 'satisfaction' is one of the vital sub-aspects, which also depends on other sub-aspects, e.g reliability, usability, understandability, etc. To satisfy user experience, it is very important that 'satisfaction' is acceptable; otherwise, user experience will be negative.

In the category of subjects 'students only' there are two studies – Colbert (2004) and Ras et al. (2007) – in which the sub-aspect 'communication' (between students) was analysed. These studies are different from one another. Colbert (2004) undertook a study of university students and their communication before and during a rendezvous. The results showed that communications may be influenced by other factors. The authors wrote, 'During rendezvous, students rated the experience of communication as more frustrating, more disruptive and less socially acceptable, than before their rendezvous'. Ras et al. (2007) used an adapted free and open-source Wiki-based system called software organization platform (SOP) to support students' observations and experiences. The results showed that this standard Wiki application could improve students' communication. Furthermore, the results confirmed that QoE was sufficient, but it is still need of more research and improvement. According to these studies, we can conclude that by using subjective measurements, the authors got different findings. The sub-aspect 'communication' is dependent on other factors such as 'disruption' and 'frustration'. To achieve a positive effect, other factors must have positive effects too.

The category 'game players' contains two studies: Wattimena et al. (2006) and Trinh et al. (2009), one being very different from the other. Wattimena et al. (2006) developed the *Quake IV* G-model and this model can enable to predict gamers' QoE of *Quake* game, based on 'jitter' and 'ping' values. The results showed that this model showed subjective data with high correlation (R=0.98) and the individual subjective opinions of game players had a satisfactory correlation (i.e. between 0.74 and 0.95). Another important finding was that higher 'ping' times negatively affected both the subjective QoE of players and the objective gaming quality. 'Jitter' had the same influence as 'ping'. The third important finding was the positive results of the sub-aspect 'packet loss', because the *Quake IV* engine is very robust.

Another study, Trinh et al. (2009), had different interests. The authors used game theory analysis for static and dynamic Internet Service Provider (ISP) markets. The results showed that socio-economic aspects, such as cost and pricing model, are important in the marketplace. Despite the fact that there were only two studies conducted, I can conclude that Real-time collaboration (i.e. gaming through the internet) requires stable connections and throughput. Furthermore, the sub-aspects 'jitter' and 'ping' are very sensitive here. I can say that subjective and objective

measurements depend on and are influenced by these sub-aspects. More research is needed to confirm the findings.

Type of study

The results (Chapter 4.2, Figure 7) showed that the most evaluated 'type of study' with human participants was 'experimental study' (48 per cent), the next most evaluated and analysed 'type of study' was 'survey' (31 per cent) and the least evaluated was 'case study' (20.8 per cent). According to these findings, we can conclude that researchers who completed studies with human participants had freedom to choose 'type of study'. For example, to complete online questionnaires (Chapter 2.4.3), which are becoming more and more popular among researchers, is cheaper and less time-consuming. Therefore, researchers have undertaken 'survey' studies (31 per cent), and not only experimental studies. Another point is that the 'case study' method, which can be used to investigate a particular problem or situation, for example, in a company (Chapter 2.4.1), is also commonly used method in the studies with human participants.

5.2 Studies without human participants

The sub-aspects 'loss / packet loss', 'jitter', 'delay', 'bandwidth' and 'quality of video' belong to empirical studies that have been conducted without human participants (Figure 17, Chapter 4.5).

As described in the chapter 4.1.2, Table 5 shows that the topic 'multimedia streaming' was of great interest to researchers. There are two studies that analysed behaviour of multimedia streaming over IP, while other studies were undertaken with multimedia streaming over a wireless network. All studies were either concentrated on objective metrics, such as Pick Signal-to-Noise Ration (PSNR), Video Quality Measurement (VQM), or techniques such as VoIP-Codec Switching. Therefore, it is difficult to draw conclusions about multimedia streams over IP or a network. There are still some issues that should be noted. The results showed that by using subjective and objective end user evaluations, the QoE could be achieved by using and iterating two layers, i.e. application QoS (AQoS) and network QoS (NQoS), and these layers were considered as a whole rather than as a single entity (Siller et al., 2003). Another issue is that VoIP-Codec Switching technique is significant and important technique that will positively contribute to mobile users' quality of experience (Waltermann et al., 2008). In another study, the behaviour of multimedia streams, which are delivered over IP networks, has been described. The experiment was described as a 'Black box' system. There are some non-controllable factors, such as packet loss rate and delay, which interact with controllable factors like video bitrates, to produce a response. The results showed that by using the statistical technique Pareto analysis, the sub-aspect 'delay' played an important role and affected the video quality QoE (Lopez et al., 2006).

The topic ‘wireless network’ has been studied by researchers, as shown in chapter 4.1.2. As it shown in Table 4, there are three studies that belong to the topic ‘wireless network’. Wireless networks such as IEEE 802.11a, IEEE 802.11g and IEEE 802.16 (WiMAX) were analysed. Both IEEE 802 11 and WiMAX were satisfactory in order to obtain higher network’s QoE parameter for IPTV in two studies (Garcia et al., 2009 and Cosma et al., 2008). Garcia et al. (2009) also specified that IEEE 802.11a was better than IEEE 802.11g, although these results should be replicated with another network because of the effect of interference using a IEEE 802.11 b/g university network. According to these findings, we can conclude that two wireless technologies IEEE 802.11 and IEEE 802.16 (WiMAX) achieve satisfaction results. Achieving a higher QoE parameter for IPTV was enabled by using these two technologies. Another point from these findings was that there is still uncertainly between two wireless technologies, IEEE 802.11a and IEEE 802.11g. Findings show that IEEE 802.11a can achieve better satisfaction than other one. To make these conclusions is difficult, because the experiment was done in one environment. To achieve more useful results, it should be tested for example with different networks. Secondly, more research should be done.

Furthermore, both studies – Garcia et al. (2009) and Cosma et al. (2008) – analysed the sub-aspect ‘bandwidth’. In the study of Garcia et al. (2009), in which effective bandwidth is described, the results showed that an IPTV content distribution system could support 5 IPTV channels in the wireless network WiMAX scenario. Cosma et al. (2008) described the same sub-aspect, and it was tested with the Universal Mobile Telecommunications System (UMTS) client. The results showed that the UMTS client had bandwidth limitations. Therefore, with the lower quality settings, QoE was satisfactory. Two studies analysed the same sub-aspect within different environments. According to these findings, it seems that WiMAX is a robust network. As mentioned earlier, by using WiMAX technology it is possible to achieve satisfactory results, but the usefulness of these findings still needs confirmation. There is still need for more research on this topic in order to confirm or disprove these results.

There are four studies that belong to the topic ‘Network traffic’ (Chapter 4.1.2, Table 4). In three of these studies, some traffic characteristics of Transmission Control Protocol (TCP) connections were analysed. Navarro-Ortiz et al. (2008) proved that TCP applications are handled efficiently and interactive services will achieve the required performance. While the results from other two studies – Collange et al. (2008) and Dyck et al. (2004) – present different results. Collange et al. (2008) found that due to the retransmission by TCP of a lost packet, there is at least one packet that will be lost in the last congestion window. Therefore, the authors proved that there is a relationship between traffic and network performance. For example, if the loss rate is high, then TCP connections should be shorter and smaller. From another point of view, Dyck et al. (2004) wrote ‘Results reiterate that TCP is unsuitable for real-time awareness data on real-world networks. Many research systems still use TCP, but this protocol is unable to meet QoE requirements in all, but the best network conditions’. It can be concluded that there is still some disagreement concerning this TCP protocol. Findings show that TCP is unsuitable for real-time

awareness data on a network. To confirm these findings, we still need more research on TCP and different aspects.

The next two topics, 'web services' and 'web search engine', were a less interesting area for researchers it seems, as there has been only one study undertaken for each. For example, Pandey et al. (2005) undertook a study that introduced new Web crawling paradigm for search engines. The motive for that was to allocate resources to crawling tasks and make it possible to achieve and maximize the quality of the user experience. It is too early to draw any conclusions from this. More research in both topics should be undertaken.

Type of Study

The results, presented in Chapter 4.2, Figure 8, showed that the most usual 'type of study' without human participants was 'experimental study' (93 per cent), while the least common was 'case study' (7 per cent). According to these findings, we can conclude that to undertake a study without human participants, the researcher chooses one dominant method, namely 'experimental study' (Chapter 2.4.2). The first reason is that this method investigates the relationship between two or more different variables in a special environment. By using an experiment, researchers have the possibility to keep these variables under observation directly and systematically. Usually, these experiments were done in laboratories. The second reason is that the researchers have one or more hypotheses and, by using different variables, they have the possibility to test them.

5.3 Summaries

This section short summarizes the findings between human and without human participants, which were described above.

The main findings from studies **with human participants:**

- QoE can be achieved when services or techniques such as IPTV or P2PTV are working properly. When the amount of users increase, the technology P2PTV will do less blocking compared to IPTV.
- 'Satisfaction' is one of the vital sub-aspects, which also depends on others sub-aspects, e.g. 'reliability', 'usability', 'understandability' etc.
- The factor 'communication', which depends on other factors such as 'disruption' and 'frustration', is also an important element when seeking to achieve better QoE.
- Real-time collaboration requires stable connection and throughput. The sub-aspects 'jitter' and 'ping' are sensitive. Subjective and objective measurements depend on and are influenced by 'jitter' and 'ping'.

- Different types of studies have been performed: ‘experimental studies’ (48%), ‘surveys’ (31%) and ‘case studies’ (20.8%).

The main findings from studies **without human participants**:

- QoE can be achieved by using and iterating two layers such as application QoS (AQoS) and network (NQoS) and these layers are considered as a whole entity.
- Both IEEE 802.11 and WiMAX were satisfactory in order to gain a higher network’s QoE parameter for IPTV.
- VoIP-Codec Switching technique is significant and important. By using this technique, a mobile user’s QoE can be achieved.
- Non-controllable factors such as ‘packet loss’ and ‘delay’ are important, as they can affect the video quality QoE.
- An IPTV content distribution system could support 5 IPTV channels in the wireless network WiMAX, but it depends on effective bandwidth.
- Replicated results show that TCP protocol is unable to meet QoE requirements.

Experimental studies have been performed most often (93%).

6. Threats to Validity

This section will discuss potential threats to the validity of this research. Creswell (2002) and Wohlin et al. (2000) have defined threats to validity as follows.

- **Internal validity**: there is internal validity if an observed relationship between the treatment and the outcome is causal and not the result of other factors, which may hinder the drawing of correct inferences from data.
- **External validity:** the extent to which the results of the study can be generalized outside the scope of the study, i.e. applied to other social or racial groups than those that were studied.
- **Construct validity** concerns the relationship between theory and observation. Threats to construct validity occur because of the use of inadequate definitions of the measured variables.

6.1 Internal validity

The best way to address this threat to validity is to involve a few independent researchers. They should extract data independently from each other and afterwards compare, discuss and analyse their findings.

Following discussion with my advisor, we selected the topic QoE and I selected and analysed 44 empirical studies. This was done by one researcher (myself) and the result of my work has not been verified or analysed by other researchers.

Another issue was to develop a framework of QoE. After I had selected all aspects and sub-aspects from each of the selected studies, I grouped them into the main categories. The grouping procedure is described in chapter 4.

6.2 External validity

Well-known researchers, i.e. B. Kitchenham, A. Fink and the University of North Carolina (UNC), have written papers and manuals on how to conduct a literature survey. I have carefully analysed them and I have written about this in chapter 3.1.1, chapter 3.1.2 and chapter 3.1.3. I decided to follow systematically Fink's written manual, which helped me to avoid bias. All these steps are described in chapter 3.2

Another issue is that the HCI database (http://hcibib.org/) is not functioning as well as other databases I have used, because it is unable to answer queries that for instance contain the word 'of'. This might be due to their choice of search engine software used to access the database.

However, to look further into this issue is outside the scope of this thesis. More explanation can be found in Appendix B.

6.3 Construct validity

The term ‘construct validity’ is not directly applicable to the present thesis. As far as I know, there is no similar research that has been done by others. I have not used hypotheses or questions to analyse all the studies, so there was no construct to validate.

In this sense, I strongly believe that my work can contribute to further research by other researchers on the field in the future.

7. Conclusion and Future work

This chapter summarizes the main findings and at the end I propose future work.

7.1 Conclusions

To conclude, I have done the following:

- Listed some definitions, models, ISO standards and empirical methods to understand the concept Quality of Experience.
- Developed a framework, which was useful to analyse all aspects and sub-aspects, which were selected from all empirical studies.
- Completed, by using the A. Fink method, a systematic literature survey, which was the main focus of the present thesis.

The main findings are:

According to the best of our knowledge, this is the first systematic literature review of this topic. A framework of QoE, based on Alben's *The ACM interaction awards criteria* model, was developed, creating the possibility of grouping together and analysing all the selected empirical studies. The framework may be easily applied and used by other researchers.

Totally, 44 empirical studies were selected and analysed from seven different electronic databases. The studies were grouped into studies with human participants (66 per cent) and studies without human participants (34 per cent). It was found that the studies with human participants consist of experimental studies, surveys and case studies, while the studies without human participants consist of experimental studies and only one case study.

The results from this literature survey indicate that the majority of researchers were most interested in aspects groups such as 'learnability and usability', 'network – QoS', 'manageable', 'aesthetic experience', 'mean opinion score' and 'needed'.

The sub-aspect 'satisfaction', which belongs to the aspect group 'learnability and usability', was one of the most evaluated sub-aspects by the researchers, often considered to be nothing less than vital. The results suggests that 'satisfaction' can also be influenced by other sub-aspects such as 'reliability', 'usability', 'understandability', 'jitter', 'ping', etc. These aspects must function well, otherwise it is impossible to gain QoE of user satisfaction. Another point is that by using learning system which was developed for students, user's satisfaction can be achieved.

The next most evaluated sub-aspect was ‘loss/packet loss’, which belongs to the aspect group ‘network – QoS’. Firstly, using the technology Wifi 802.11, the QoE was excelent on both audio and video transmissions. Secondly, ‘loss/packet loss’ dependend on the other sub-aspects ‘bandwidth’ and ‘response time’.

The findings from this study show that different sub-aspects depend on different services, for example, wireless technology or multimedia streaming. It is not enough that one sub-aspect functions very well, because most of the sub-aspects are closely related to each other. Therefore, it is very important that sub-aspects, which are dependent on each other, are functioning as one group to achieve higher QoE of user experience.

The results from this literature survey indicate that there is little research done on sub-aspects like ‘understandability’, ‘navigation’, ‘zooming’, etc. According to mobile technology, ‘zooming’ is a vital sub-aspect, because mobile technology is developing at a very high speed and mobile phones have become more and more popular. Therefore, researchers should give more attention to this sub-aspect.

The largest amount of human subjects participated in one study belonging to the category ‘TV channels’, but it is the only such study performed. To gain more knowledge in this field, researchers should perform more studies.

There is little existing knowledge about topics such as ‘web services’ and ‘web search engine’ . These topics were unfortunately of little interest to researchers and should be adressed in future research.

Finally, QoE is a relatively new phenomenon which still needs time to mature. I believe that I have presented relevant results on the phenomenon. I hope this thesis will help researchers to gain a better understanding of this concept.

7.2 Future work

This section is an overview of possible future research on the phenomenon Quality of Experience. The areas, which require further investigation, will be discussed. These areas may be relevant for researchers.

Some aspects have been studied only once. Therefore, researchers should pay more attention to aspects like ‘social economic’, ‘understandability’, ‘availability’, ‘navigation’ and ‘zooming’, etc. At the same time more research on topics such as ‘web services’, ‘web search engine’ or the category ‘TV channels’ is needed.

Future work that could elaborate on the results of this research, would be a study of more articles in the same year's period. It was described earlier in the present work how relevant studies were selected (Chapter 3.2.5). New research could validate our findings and at the same time give results that are more correct.

To replicate the same study, it would be of great interest to involve a number of researchers, to eliminate all possibilities of bias or doubt, and to confirm or disprove results.

Finally, it would be of great interest to see how QoE evolves in the market in the future, i.e. from 2009 to 2012. The results could give us more understanding of how researchers will comprehend the new phenomenon QoE in their work.

List of Acronyms and Abbreviations

ADSL	Asymmetrical Digital Subscriber Line
AHS	Adaptive Hypermedia Systems
ANOVA	Analyses of variance
APs	Access Point
AQoS	Application Quality of Service Layer
DSLAM	Digital Subscriber Line Access Multiplexers
EBR	Embarrassment-based Refreshing
ETSI	European Telecommunications Standards Institute
GUI	Graphical User Interface
HCI	Human-Computer Interaction
HPT	High-Performance Telepointers
ICT	Information and Communication Technologies
IEC	International Electro technical Commission
IEEE	The Institute of Electrical and Electronics Engineers
IPDV	Inter-Packet-Delay Variation
IPTV	Internet Protocol Television
ISO	International Organization for Standardization
ISP	Internet Service Provider
IT	Information Technology
LDA+	Loss-Delay Adaption Algorithm
MOS	Mean Opinion Score
MPLS	Multiprotocol label switching
NGI	Next Generation Internet
NQoS	Network Quality of Service
OSI Model	Open Systems Interconnection Reference Model
OWD	One-Way Delay
PESQ	Perceptual evaluation of speech quality
PLR	Packet Loss Ratio
PSNR	Pick Signal-to-Noise Ration
PSQA	Pseudo-Subjective Quality Assessment
P2P	Peer-to-peer
P2PTV	Peer-to-peer TV software application
QOAS	Quality-Oriented Adaptive Scheme
QoE	Quality of Experience
QoEAHS	QoE Adaptive Hypermedia System

QoS	Quality of Service
ROIAS	Region of Interest-based Adaptive Scheme
R2D2 Networks	Road to media-aware user-Dependent self-aDaptive Networks
SBR	Staleness Based Refreshing
SLR	Systematic literature review
SOP	Software organization platform
SQ	Software Quality
TCP	Transmission Control Protocol
TFRC	TCP-Friendly Rate Control Protocol
UDP	User Datagram Protocol
UMTS	Universal Mobile Telecommunications System
UNC	University of North Carolina
URL	Uniform Resource Locator
VoIP	Voice over IP
VQM	Video Quality Measurement
WiMAX	Worldwide Interoperability for Microwave Access

References A

Abran, A., Suryn, W., Khelifi, A., Rilling, J., Seffah, A., & Robert, F. (2003). Consolidating the ISO Usability Models. *Paper presented at 11th annual International Software Quality Management Conference*.

Alben, L., & Faris, J. (2003). Criteria for interaction design, from http://www.albendesign.com/albenfaris/images/criteria.html [Visited: 10.10.2009]

Alben, Lauralee (1996). Quality of experience: defining the criteria for effective interaction design. *interactions, 3*(3), 11-15.

Aldrich, S. E., Marks, R. T., Frey, M. M., Goulde, M. A., Lewis, J. M., & Syeybold, P. B. (2000). What kind of the Tottal Customer Experience Does Your E-business Deliver? *Patricia Seybold Group*.

Alvestrand (2009). ITU-T., from http://en.wikipedia.org/w/index.php?title=ITU-T&oldid=328663416 [Visited: 30.11.2009]

Andersen, I. (1990). *Valg af Organisations-Sociologiske Metoder. Et kombinationsperspektiv. 1.*: Auflage. København.

Bauer, B., & Patrick, A. S. (2004). A Human Factors Extension to the Seven-Layer OSI Reference Model, from http://www.andrewpatrick.ca/OSI/10layer.html [Visited: 11.11.2009]

Berander, P., Damm, L., Eriksson, J., Gorschek, T., Henningsson, K., Jonsson, P., et al. (2005). *Software quality attributes and trade-offs.* Unpublished PhD, Blekinge Institute of Technology.

Bevan, N. (2009). ISO 9241: Ergonomic requirements for office work with visual display terminals from http://www.usabilitynet.org/tools/r_international.htm#9241-1x [Visited: 3.08.2009]

Boehm, B. W., Brown, J. R., & Lipow, M. (1976). *Quantitative evaluation of software quality*. Paper presented at the Proceedings of the 2nd international conference on Software engineering.

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2006). Lessons from applying the systematic literature review process within the software engineering domain. *J. Syst. Softw., 80*(4), 571-583.

Creswell, J. W. (Ed.). (2003). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*: Sage Publications, Inc.

Crosby, P. B. (1979). *Quality is free. The Art of Making Quality Certain*: McGraw-Hill Book Company.

Czaja, R., & Blair, J. (Eds.). (2005). *Designing surveys. A Guide to Decisions and Procedures*: Sage Publications, Inc.

Dittrich, Y., John, M., Singer, J., & Tessem, B. (2007). Editorial: For the Special issue on Qualitative Software Engineering Research. *Inf. Softw. Technol., 49*(6), 531-539.

Fink, A. (1995). *How to Analyze Survey Data (Survey Kit, 8)* (Vol. 8): Sage Publications, Inc.

Fink, A. (2009). *Conducting Research Literature Reviews: From the Internet to Paper* (3rd ed.): Sage Publication, Inc.

Fisher, R. A. (2009). Empirical Research, from http://www.answers.com/topic/empirical-research-2 [Visited: 20.11.2009]

Garvin, D. A. (1984). What Does 'Product Quality' Really Mean? *Sloan Management Review, 1*, 25-48.

Goodchild, J. (2005). "Integrating data voice and video - Part II," IP Video Implementation and planning guide,. *United States Telecom Association*.

Harning, M. B., & Vanderdonckt, J. (2003). CLOSING THE GAPS: Software Engineering and Human-Computer Interaction.

Hassenzahl, M., Beau, A., & Burmester, M. (2001). Engineering joy. *IEEE Software* 2-8.

ISO (2009). ISO 9126, from http://www.cis.gsu.edu/~ghubona/cis8300/ISO9126.pdf [Visited: 2.09.2009]

Jain, R. (2004). Quality of Experience. *IEEE Multimedia, Vol. 11*(1), 96-97.

Johnson, T. (2009). ITU-T in brief, from http://www.itu.int/net/ITU-T/info/Default.aspx [Visited: 21.09.2009]

Kitchenham, B. (2004). *Procedures for Performing Systematic Reviews*. Australia: Joint Technical Report Software Engineering Group, Department of Computer Science Keele University, United King and Empirical Software Engineering. .

Kitchenham, B., Brereton, P., Budgen, D., & Turner, M. (2007). 2nd International workshop on realising evidence-based software engineering (REBSE-2): overview and introduction. [Conference Paper].

Kitchenham, B., Brereton, P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering - A systematic literature review. [Journal Paper]. *Information and Software Technology, 51*(1), 7-15.

Kitchenham, B., & Pfleeger, S. L. (1996). *Software quality: the elusive target [special issues section]*. *Software, IEEE, 13*(1), 12-21.

Lopez, D., Gonzalez, F., Bellido, L., & Alonso, A. (2006). Adaptive multimedia streaming over IP based on customer oriented metrics. [Conference Paper]. *Proceedings of ISCN'06 7th International Symposium on Computer Networks*.

McCall, J. A., Richards, P. K., & Walters, G. F. (1977). Factors in software quality. Volume I. Concepts and definitions of software quality. *General Electric CO Sunnyvale California, Technical Report ADA049014*.

McInnes, A. (2009). ISO/IEC 9126, from http://en.wikipedia.org/w/index.php?title=ISO/IEC_9126&oldid=68013597 [Visited: 7.08.2009]

Neill, J. (2007). Qualitative versus Quantitative Research: Key Points in a Classic Debate, from http://wilderdom.com/research/QualitativeVersusQuantitativeResearch.html [Visited: 3.12.2009]

Nielsen, J. (1993). *Usability Engineering*: Academic Press, Inc.

O'Neil, T. M. (2002). *Quality of Experience and Quality of Service for IP Video Conferencing*: White Paper by Polycom.

Patrick, A. S., Singer, J., Corrie, B., Noel, S., El, Khatib K., Emond, B., et al. (2004). A QoE sensitive architecture for advanced collaborative environments. [Conference Paper]. *Proceedings. First International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks. IEEE Comput. Soc.*

Preece, J., Rogers, Y., & Sharp, H. (2006). *INTERACTION DESIGN: Beyond Human-Computer Interaction* (2nd ed.): Wiley & Sons.

Renier, M. (2009). Software quality, from http://en.wikipedia.org/w/index.php?title=Software_quality&oldid=315002355 [Visited: 20.9.2009]

Roldan, L. M. P. (2008). *Road to media-aware user-Dependent self-aDaptive NETWORKS*. Madrid, Spain: Telefonica I+D.

Rubino, G. (2005). Quantifying the quality of audio and video transmissions over the internet: the psqa approach. *In J.Barria, editor, Design and Operations of Communication Networks: A review of Wired and Wireless Modelling and Management Challenges.*

Rubino, G. (2009). PSQA: Pseudo-Subjective Quality Assesment, from http://ralyx.inrialpes.fr/2004/Raweb/armor/uid34.html [Visited: 13.11.2009]

Shewhart, W. A. (1931). *Economic Control of Quality of Manufactured Product.*: New York: D. Van Nostrand Co.

Siller, M., & Woods, J. (2003). Improving quality of experience for multimedia services by QoS arbitration on a QoE framework. Packet Video. *Nantes*.

Thackara, J. (2001). Why is interaction design important? , from http://www.doorsofperception.com/archives/2001/11/why_is_interact.php [Visited: 23.11.2009]

UNC Press (2009). Literature Reviews, from http://www.unc.edu/depts/wcweb/handouts/literature_review.html [Visited: 10.10.2009]

Unknown (2000). Seriec G: Transmission Systems and Media, Digital Systems and Networks: International telephone connections and circuits - General definitions.

Unknown (2009a). Malcolm Baldrige National Quality Award, from http://www.nist.gov/public_affairs/factsheet/mbnqa.htm [Visited: 09.10.2009]

Unknown (2009b). Quality of Experience, from http://searchcrm.techtarget.com/definition/Quality-of-Experience [Visited: 09.09.2009]

Unknown (2009c). Quality of Service, from http://technet.microsoft.com/en-us/network/bb530836.aspx [Visited: 14.01.2009]

Unknown (2009d). Recommendation ITU-T P.862: Perceptual evaluation of speach quality (PESQ), from http://www.itu.int/net/itu-t/sigdb/genaudio/Pseries.htm [Visited: 13.11.2009]

Unknown (2009e). *SearchCIO-Midmarket.com Definitions*, from http://searchcio-midmarket.techtarget.com/sDefinition/0,,sid183_gci928405,00.html [Visited: 21.09.2009]

Unknown (2009f). User Interface design, from http://www.use-design.com/eng/design/user_interface_design.php [Visited: 5.10.2009]

Van Moorsel, A. (2001). Metrics for the Internet Age: Quality of Experience and Quality of Business. *The Fifth Performability Workshop*.

Weinberg, G. (1991). Quality Software Management - Systems Thinking (Vol. 1). New York: Dorset House Publishing.

Wohlin, C., Runeson, P., Host, M., Ohlsson, M. C., Regnell, B., & Wesslen, A. (2000). *Experimentation in software engineering: an introduction*: Kluwer Academic Publishers.

Xiao, X., & Ni, L. M. (1999). Internet QoS: The Big Picture. *IEEE Network Magazine*, 8-18.

Yeh, Y. W. (2009). Information and communication technologies, from http://en.wikipedia.org/w/index.php?title=Information_and_communication_technologies&oldid=335821589 [Visited: 5.01.2010]

Yin, Robert K. (2003). *Case study research - Design and Methods* (Vol. 5): SAGE Publications.

Zapater, M. N., & Bressan, G. (2007). A proposed approach for quality of experience assurance for IPTV. [Conference Paper]. *First International Conference on the Digital Society*.

Zekone (2009). Experiment, from http://en.wikipedia.org/w/index.php?title=Experiment&oldid=313952431 [Visited: 15.09.2009]

Ziegler, P. (2004). *Department of Information Technology Glossary*. from http://www.nh.gov/doit/internet/org/cio/hr/forms/documents/OITGlossary.pdf [Visited: 13.08.2009].

References B

[1] Cheung, C. M. K., & Lee, M. K. O. (2005). *The Asymmetric Effect of Website Attribute Performance on Satisfaction: An Empirical Study.* Paper presented at the System Sciences, 2005. HICSS '05. Proceedings of the 38th Annual Hawaii International Conference on.

[2] Mahlke, S. (2002). *Factors influencing the experience of website usage*. Paper presented at the CHI '02 extended abstracts on Human factors in computing systems.

[3] Muntean, C. H., & McManis, J. (2006). The value of QoE-based adaptation approach in educational hypermedia: empirical evaluation. [Conference Paper]. *Adaptive Hypermedia and Adaptive Web Based Systems. 4th International Conference, AH, 4018*, 121-130.

[4] Knoche, H., Papaleo, M., Sasse, M. A., & Vanelli-Coralli, A. (2007). *The kindest cut: enhancing the user experience of mobile tv through adequate zooming*. Paper presented at the Proceedings of the 15th international conference on Multimedia.

[5] Colbert, M. (2004). User experience of communication before and during rendezvous: interim results. [Journal Paper]. *Personal and Ubiquitous Computing, 9*(3), 134-141.

[6] Bhattacharjee, S., Gopal, R. D., Lertwachara, K., & Marsden, J. R. (2003). *Economic of online music*. Paper presented at the Proceedings of the 5th international conference on Electronic commerce.

[7] Wattimena, A. F., Kooij, R. E., Vugt, J. M. van, & Ahmed, O. K. (2006). *Predicting the perceived quality of a first person shooter: the Quake IV G-model*. Paper presented at the Proceedings of 5th ACM SIGCOMM workshop on Network and system support for games.

[8] Van Velsen, L. S., Steehouder, M. F., & de, Jong M. D. T. (2007). Evaluation of user support: factors that affect user satisfaction with helpdesks and helplines. [Journal Paper]. *IEEE Transactions on Professional Communication, 50*(3), 219-231.

[9] Siller, M., & Woods, J. C. (2003). *QoS arbitration for improving the QoE in multimedia transmission.* Paper presented at the Visual Information Engineering, 2003. VIE 2003. International Conference on.

[10] Ras, E., Carbon, R., Decker, B., & Rech, J. (2007). Experience management Wikis for reflective practice in software capstone projects. [Journal Paper]. *IEEE Transactions on Education, 50*(4), 312-320.

[11] Ali, A. S., Ludwig, S. A., & Rana, O. F. (2005). A cognitive trust-based approach for Web service discovery and selection. [Conference Paper]. *Proceedings. Third European Conference on Web Services. IEEE Computer Society, 12*.

[12] Waltermann, M., Lewcio, B., Vidales, P., & Moller, S. (2008). A technique for seamless VoIP-Codec switching in next generation networks. [Conference Paper].

[13] Muntean, G. M., & Cranley, N. (2007). Resource Efficient Quality-Oriented Wireless Broadcasting of Adaptive Multimedia Content. *IEEE Transactions on Broadcasting, 53(1)*, 362-368.

[14] Silzle, A. (2007). *Generation of Quality Taxonomies for Auditory Virtual Environments by Means of a Systematic Expert Survey*. Paper presented at the 19th INTERNATIONAL CONGRESS ON ACOUSTICS MADRID, 2-7 SEPTEMBER. from http://www.sea-acustica.es/WEB_ICA_07/fchrs/papers/ppa-02-009.pdf

[15] Hamam, A., Eid, M., El Saddik, A., & Georganas, N. D. (2008). *A quality of experience model for haptic user interfaces*. Paper presented at the Proceedings of the 2008 Ambi-Sys workshop on Haptic user interfaces in ambient media systems.

[16] Lu, Yue, Kuipers, F., Janic, M., & Van Mieghem, P. (2008). E2E Blocking Probability of IPTV and P2PTV *NETWORKING 2008 Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet* (pp. 445-456).

[17] Zhang, T., Rau, Pei-Luen, & Salvendy, G. (2007). Developing Instrument for Handset Usability Evaluation: A Survey Study *Human-Computer Interaction. Interaction Design and Usability* (pp. 662-671).

[18] Collange, D., & Costeux, J.L. (2008). Passive Estimation of Quality of Experience. *Journal of Universal Computer Science, vol. 14, no. 5*, 625-641.

[19] Agrawal, D., Beigi, M. S., Bisdikian, C., & Kang-Won, Lee (2007). Planning and managing the IPTV service deployment. [Conference Paper].

[20] Monteiro, Janio M., & Nunes, Mario S. (2007, 1-4 July 2007). *A Subjective Quality Estimation Tool for the Evaluation of Video Communication Systems.* Paper presented at the Computers and Communications, 2007. ISCC 2007. 12th IEEE Symposium on.

[21] Lopez, D., Gonzalez, F., Bellido, L., & Alonso, A. (2006). Adaptive multimedia streaming over IP based on customer oriented metrics. [Conference Paper]. *Proceedings of ISCN'06 7th International Symposium on Computer Networks*.

[22] Holone, H., Misund, G., Tolsby, H., & Kristoffersen, S. (2008). *Aspects of personal navigation with collaborative user feedback.* Paper presented at the Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges.

[23] Vegiris, C.E., Avdelidis, K.A., Dimoulas, C.A., & Papanikolau, G.V. (2008). Live Broadcasting of High Definition Audiovisual Content Using HDTV over Broadband IP Networks. *International Journal of Digital Multimedia Broadcasting*, 18 pages.

[24] Ciubotaru, B., Muntean, G. M., & Ghinea, G. (2009). Objective assessment of region of interest-aware adaptive multimedia streaming quality. [Journal Paper]. *IEEE Transactions on Broadcasting, 55*(2), 202-212.

[25] Dyck, J., Gutwin, C., Subramanian, S., & Fedak, C. (2004). High-performance telepointers. [Conference Paper]. *Computer Supported Cooperative Work Conference Proceedings. ACM.*

[26] Calyam, P., Ekici, E., Lee, Chang-Gun, Haffner, M., & Howes, N. (2007). A "GAP-Model" based Framework for Online VVoIP QoE Measurement.

[27] Siller, M., & Woods, J. (2006). Using an agent based platform to map quality of service to experience in conventional and active networks. *IEE Proceedings - Communications, 153*(6), 828-840.

[28] Liu, J., & Issarny, V. (2007). *An incentive compatible reputation mechanism for ubiquitous computing environments*. Paper presented at the Proceedings of the 2006 International Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services.

[29] Calyam, P., Kalash, A., Gopalan, R., Gopalan, S., & Krishnamurthy, A. (2008). RICE: a reliable and efficient Remote Instrumentation Collaboration Environment. [Journal Paper]. *Advances in Multimedia, 615186*(17).

[30] Zhang, L., Gillies, M., Dhaliwal, K., Gower, A., Robertson, D., & Crabtree, B. (2009). E-Drama: Facilitating Online Role-play using an AI Actor and Emotionally Expressive Characters. *International Journal of Artificial Intelligence in Education, 19 (1)*, pp. 5-38.

[31] Garcia, M. J., Lloret, J., Edo, M., & Lacuesta, R. (2009). *IPTV distribution network access system using WiMAX and WLAN technologies*. Paper presented at the Proceedings of the 4th edition of the UPGRADE-CN workshop on Use of P2P, GRID and agents for the development of content networks.

[32] Moid, A., & Fapojuwo, A. O. (2009). A cross-layer framework for efficient streaming of H.264 video over IEEE 802.11 networks. *J. Comp. Sys., Netw., and Comm., 2009*, 1-13.

[33] Dufour, A., & Trajkovic, L. (2006). *Improving Gnutella network performance using synthetic coordinates*. Paper presented at the Proceedings of the 3rd international conference on Quality of service in heterogeneous wired/wireless networks.

[34] Navarro-Ortiz, J., Ameigeiras, P., Ramos-Munoz, J. J., & Lopez-Soler, J. M. (2008). An 802.11e HCCA scheduler with an end-to-end quality aware territory method. [Journal Paper]. *Computer Communications, 32*(11), 1281-1297.

[35] Muntean, C. H. (2007). Improving learner quality of experience by content adaptation based on network conditions. [Journal Paper]. *Computers in Human Behavior, 24*(4), 1452-1472.

[36] Schaik, P. van, & Ling, J. (2003). Using on-line surveys to measure three key constructs of the quality of human-computer interaction in Web sites: psychometric properties and implications. [Journal Paper]. *International Journal of Human-Computer Studies, 59*(5), 545-567.

[37] Morris, M. G., & Turner, J. M. (2001). Assessing users' subjective quality of experience with the World Wide Web: an exploratory examination of temporal changes in technology acceptance. [Journal Paper]. *International Journal of Human-Computer Studies, 54*(6), 877-901.

[38] Trinh, T. A. , & Gyarmati, L. (2009). Revisiting Internet access pricing for loyal customers: the long-term interaction case. [Conference Paper]. *8*, 1-3.

[39] Cosma, R., Cabellos-Aparicio, A., Domenech-Benlloch, J., Gimenez-Guzman, J., Martinez-Bauset, J., Cristian, M., et al. (2008). Measurement-based analysis of the performance of several wireless technologies. [Conference Paper].

[40] Yoo, Y., Tussyadiah, I. P., Fesenmaier, D. R., Saari, T., & Tjstheim, I. (2008). Emergent distributed narratives in spatiotemporal mobility: an exploratory study on mobile 2.0 services. [Conference Paper].

[41] Koivisto, M. (2007). *Acceptance and use of mobile phone mediated communication technologies in community communications.* Paper presented at the Digital Information Management, 2007. ICDIM '07. 2nd International Conference on.

[42] Pianesi, F., Graziola, I., Zancanaro, M., & Goren-Bar, D. (2009). The motivational and control structure underlying the acceptance of adaptive museum guides - an empirical study. [Journal Paper]. *Interacting with Computers, 21*(3), 186-200.

[43] Knoche, H., & Sasse, M. A. (2009). The big picture on small screens delivering acceptable video quality in mobile TV. *ACM Trans. Multimedia Comput. Commun. Appl., 5*(3), 1-27.

[44] Pandey, S., & Olston, C. (2005). *User-centric Web crawling*. Paper presented at the Proceedings of the 14th international conference on World Wide Web.

Appendix A

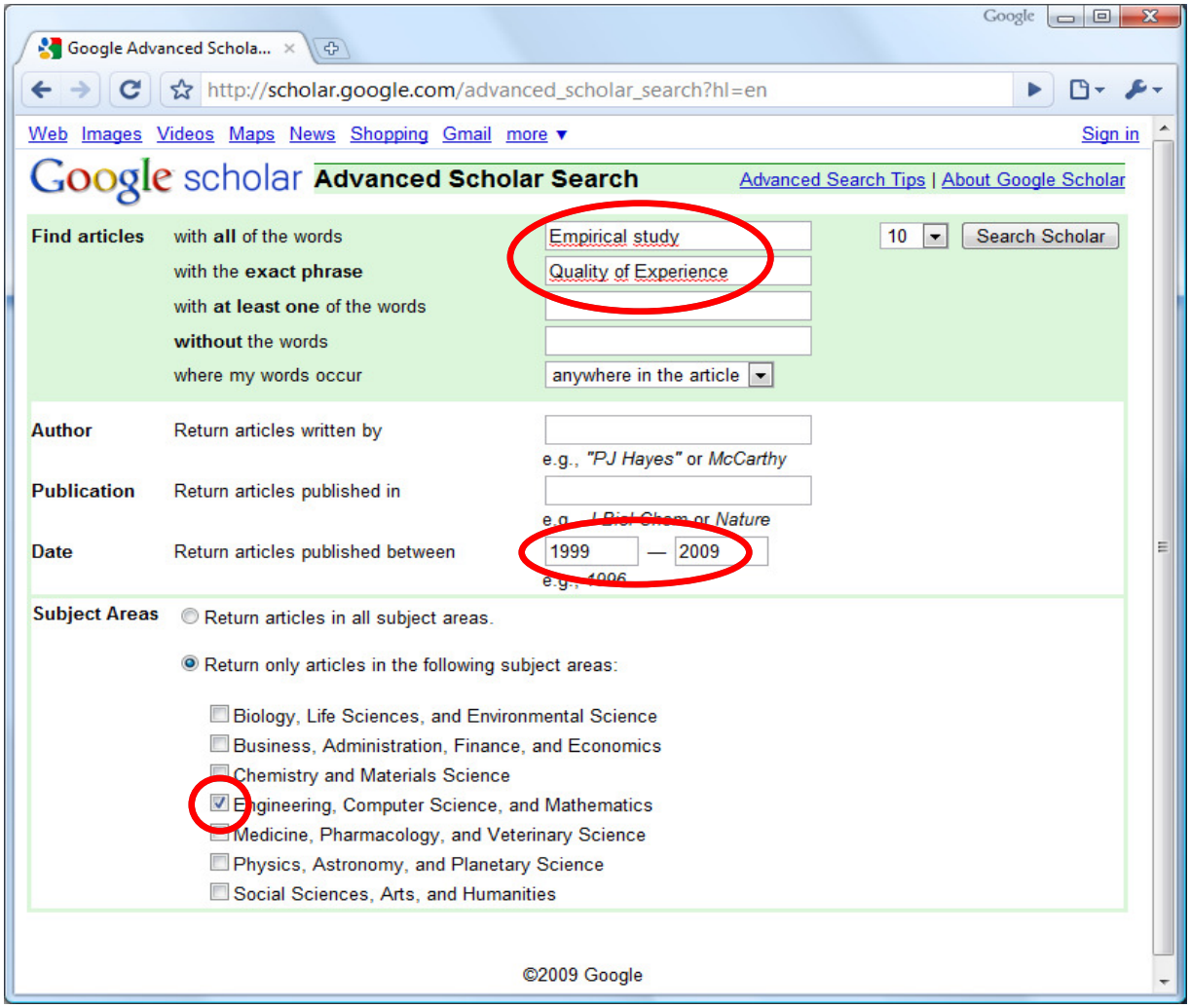


Figure 18: Search criteria in the Google Scholar database.

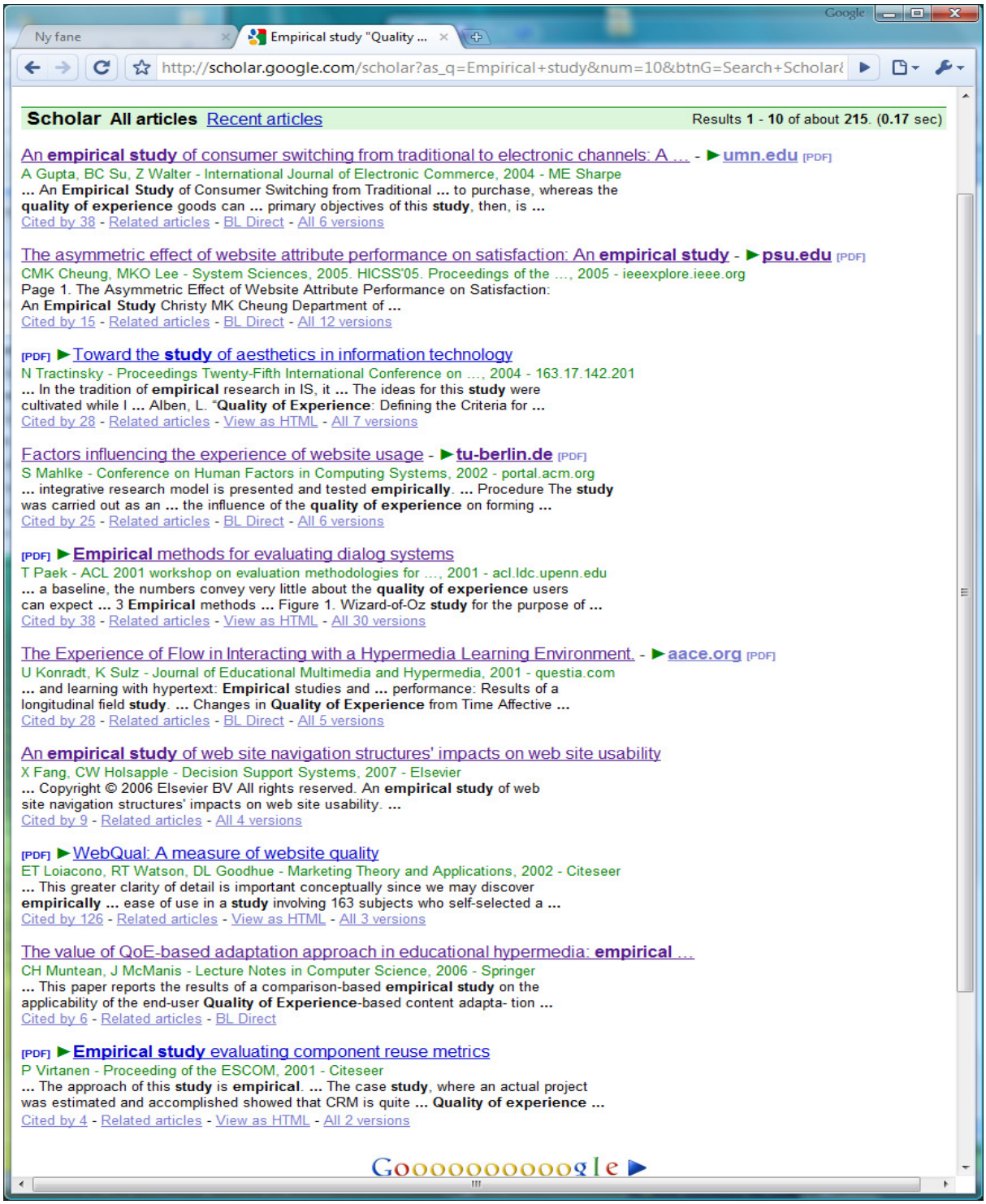


Figure 19: Results in the electronic database Google Scholar.

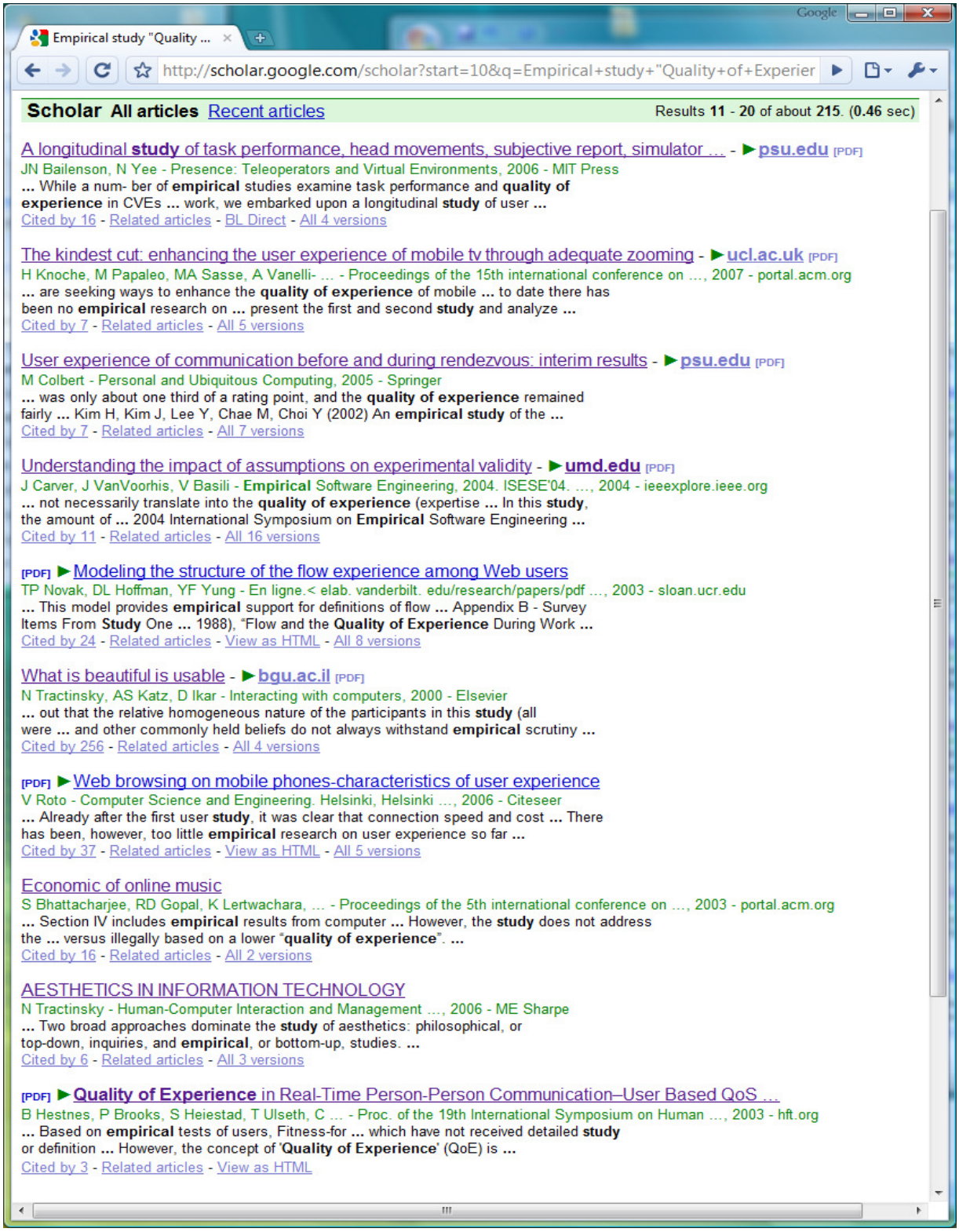


Figure 20: Results in the electronic database Google Scholar.

Appendix B - Search Engines in different databases

I had a little experiment and I tried to find out differences between two almost identical search phrases without and with double-quote (“”). The research string was defined as follows.

- Quality of Experience (string **A**)
- ‘Quality of Experience’ (string **B**)

The list of the results in Table 9.

Name of databases	String **A**	String **B**
Google Scholar	788,000	1,970
ACM Digital Library	45,778	325
IEEE Xplore	193	173
Science Direct	501,931	641
Springer Link	160,177	401
ISI Web of Knowledge	31,676	361
HCI	**199**	**199**
Total:	**1`527,954**	**4,070**

Table 9: Results **with** and **without** double-quotes in different electronic databases

Table 9 shows that the search phrase with double-quotes give fewer returned results than without quotes in almost all databases used in this experiment, with the exception of the HCI database which returned the exact same number of results to the two search phrases.

To illustrate the special case with the HCI database, see the two screenshots below (figure 21 and figure 22).

Firstly, I have used search string (String A), which was explained above. After I performed the search, I noticed that the search engine changed the search phrase from Quality of Experience into Quality & Experience (Figure 21). This show that the word “of” was ignored and that the search engine inserted a Boolean search operator “AND” (&) between the words in the search phrase. This is clearly evidence that search engine for HCI database is working in different way than other databases I have used.

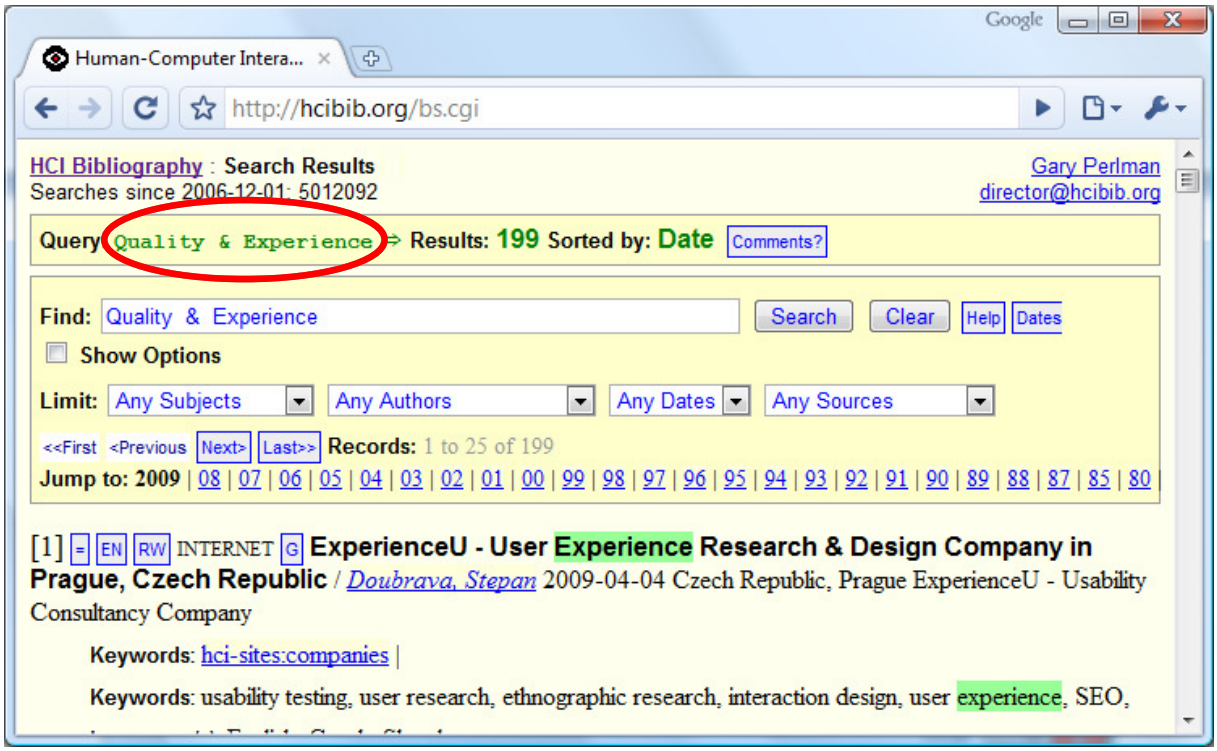


Figure 21: Results with search criteria **without** double-quotes QoE in HCI Bibliography.

Another search phrase I have used is **String B** (Figure 22). The search engine rewrote String B with the same result as for String A. Double-quotes are ignored, a warning on the search result page informs me. The word 'of' was removed as well. The conclusion is that researchers are unable to specify a phrase as search input for this search engine, since as a group of single words the search phrase is processed.

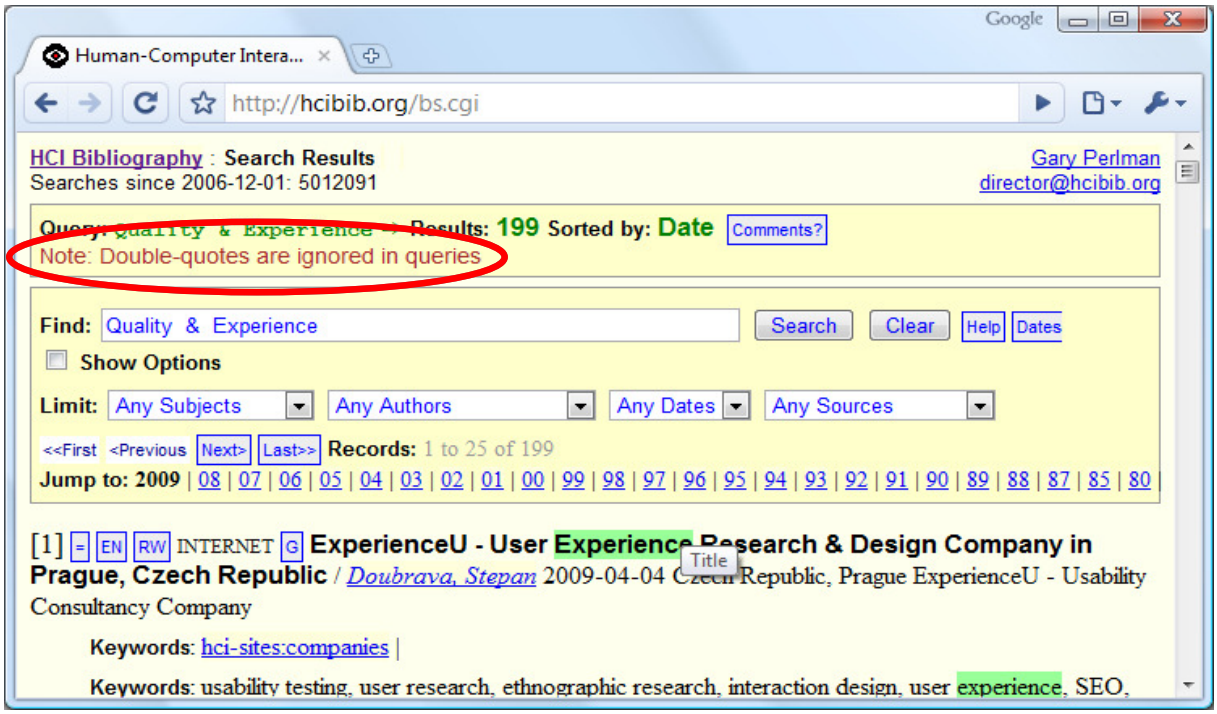


Figure 22: Result with search criteria **with** double-quotes 'Quality of Experience' in HCI Bibliography.

Appendix C - Different electronic databases

Table 10 consists of the following series of columns.

The column *Num.of hits* shows how many hits have given the searched results in different electronic databases
The column *Database* indicates the names of the electronic databases.
The column *URL* shows the Uniform Resource Locator, i.e. the web address that each databasehas.
The column % (per cent) shows how many hits (resuls) were found from each database in per cent.

Num. of hits	Database	URL	%
215	Google Scholar	http://scholar.google.com	68,5
32	ACM Digital Library	http://portal.acm.org	10,2
28	IEEE Xplore	http://ieeexplore.ieee.org	8,9
22	Science Direct	http://www.sciencedirect.com	7,0
7	Springer Link	http://www.springerlink.com	2,2
6	ISI Web of knowledge	http://apps.isiknowledge.com	1,9
4	HCI	http://hcibib.org	1,3
Total: 314			**100**

Table 10: Overview of all the databases that were used in this thesis.

Appendix D – Definitions/Explanations

Table 11 presents two main columns:

The column **name** gives the name of the concept..
The column **description** gives an explanation of each concept.

Name	Description
ITU-T	ITU-T stands for the Telecommunication Standardization Sector (**ITU-T**), which coordinates standards of telecommunications for the International Telecommunication Union (ITU). ITU-T defines elements in Information and Communication Technology (ICT) infrastructure. Communication such as voice, data or video messages between the sender and the receiver might take place with standard linking. This standard is based in Geneva, Switzerland (Alvestrand, 2009).
ITU-T Rec.P.862	Recommendation **ITU-T P.862**: Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs. It also describes a method for predicting the subjective quality of 3.1kHz (narrow-band) handsets and narrow-band speech codecs. This Recommendation gives good advice on how to use this method, presents a high-level description of that method and has shown achievable results since 1999 (Unknown, 2009d).
E-model (ITU-T Rec.G.107)	This ITU-T Recommendation gives the algorithm for the so-called E-model as the common ITU-T Transmission Rating Model. The E-model is based on the equipment impairment factor method, following previous transmission rating models. It was developed by a European Telecommunications Standards Institute (ETSI), an ad hoc group called 'Voice Transmission Quality from Mouth to Ear' (Unknown, 2000).
Pseudo-Subjective Quality Assessment (PSQA)	PSQA is a methodology and it allows one to quantify the quality of a video, audio or multimedia stream at the receiving end. This stream has passed through a packet network for example the Internet(Rubino, 2009).
ITU-T Rec.G.722	Recommendation G.722: 7kHz AUDIO-CODING within 64kBit/s.

Table 11: Definitions and descriptions

Appendix E – Description of the subjects

Table 12 presents four main columns:

The column **category of subjects** shows different groups of participants (e.g. undergraduates, graduates, professionals, etc.) in their experiments.
The column **reported subject types** gives more explanation about each category (e.g. undergraduate students, PhD students, etc.).
The column **n** shows how many studies were performed per each category
The column **% (per cent)** shows per cent of studies per group.

Category of subjects	Reported subject types	N	%
TV channel	Using the most often watched 23 TV channels. There were 297,600 TV-watchers.	1	3,4
Secondary school	Swanshurst School, a Specialist Science College in Billesley, Birmingham.	1	3,4
Undergraduates	Undergraduates, Bachelors, First-year, full-time or part-time students.	5	17,3
Graduates	PhD students.	2	6,9
Students, type unknown	Students in computer science, mathematics courses, electrical engineering, economics or undergraduate psychology students.	2	6,9
Professionals/Experts	An expert user is one who has considerable business-quality video conferencing experience due to regular usage and has in-depth system understanding. As described in this three-part categorization of users: (i) **Expert** user is one who has considerable business-quality video-conferencing experience due to regular usage and has in-depth system understanding. (ii) **General** user is one who has moderate experience due to regular usage and has in-depth system understanding (iii) **Novice** user is one who has little prior business quality videoconferencing experience but has basic system understanding .	1	3,4
Mixed group of subjects	There are six categories of internet experience: 1) No experience 2) 0 – 6 months 3) 6 months – 1 year 4) 1 – 2 years 5) 2 – 3 years 6) Over 3 years Other group descriptions: a) **Novice**: The novice users performed a set of sequential tasks with simplistic actions. b) **Expert**: The expert users performed a set of sequential tasks with advanced actions that required relatively more effort and skills.	3	10,4

Game players	Experienced internet game players.	2	6,9
Others	Tourist, individuals, online-users, theatre customers, museum visitors.	12	41,5
Total		**29**	**100**

Table 12: Overview and description of the human participants**.**

Appendix F – Overview of figures and tables

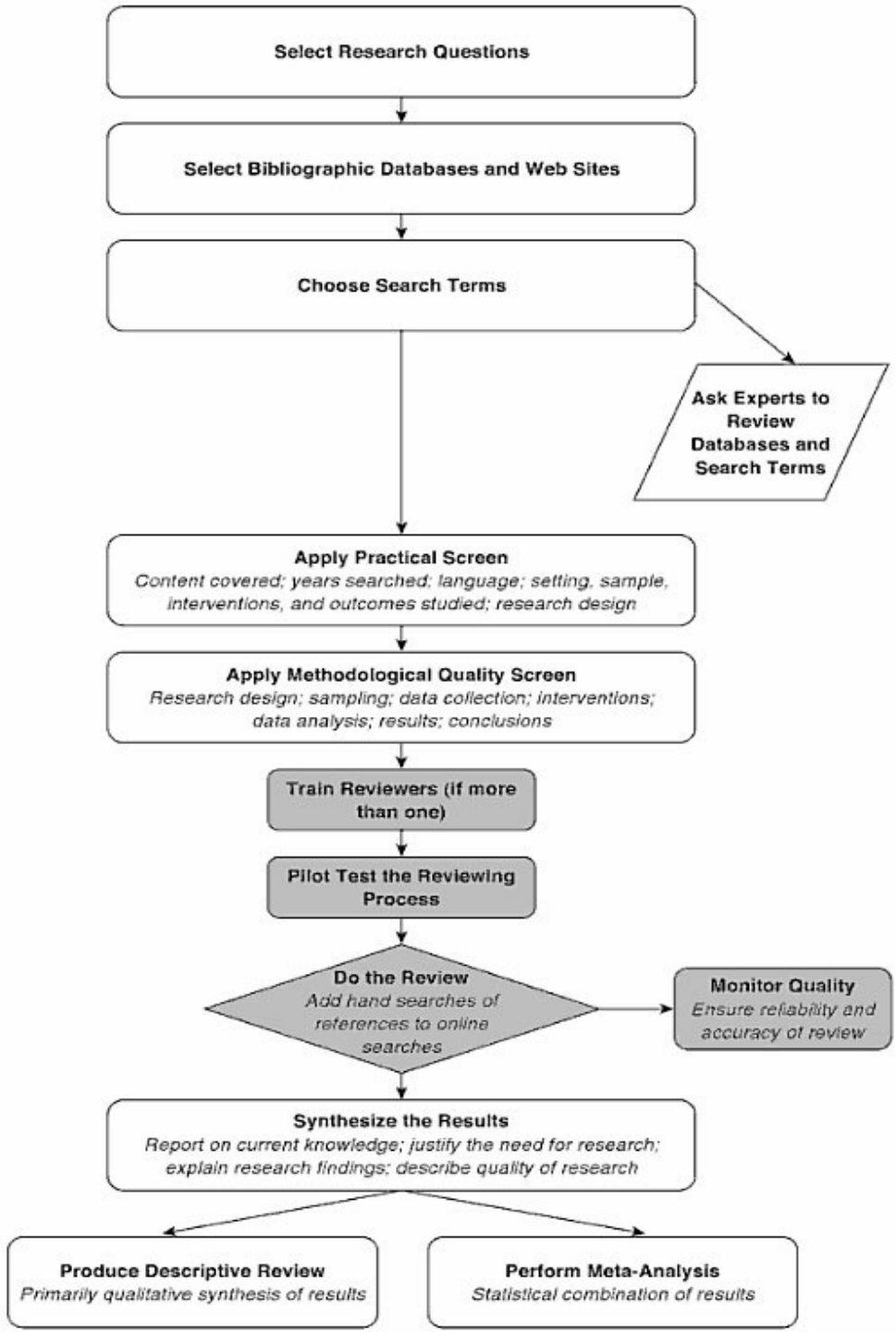


Figure 23: Steps Involved in Conducting a Research Literature Review (Fink, 2009 s.4).

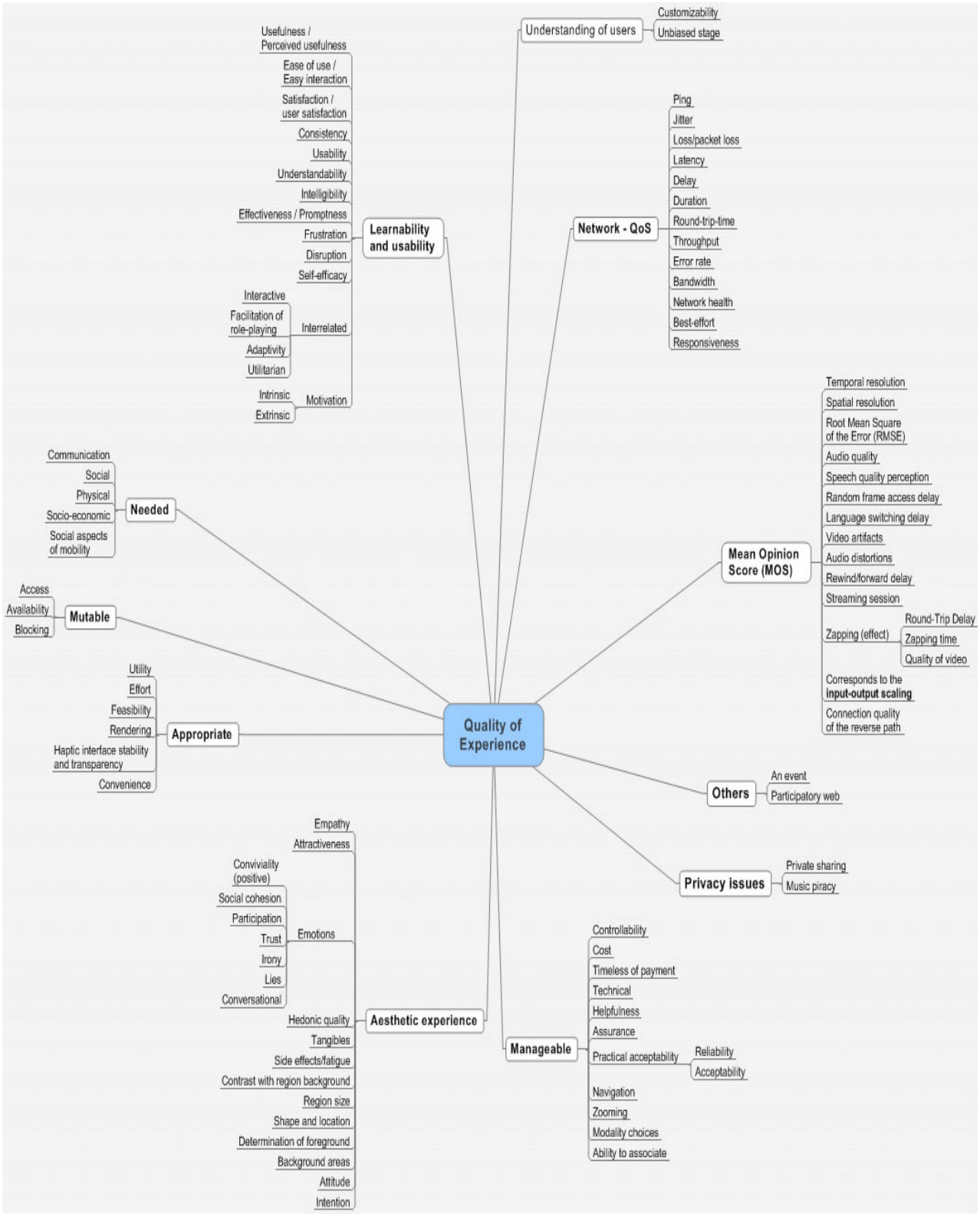


Figure 24: An overview of a framework with all aspects and sub-aspects in QoE research.

Table 13 presents all aspects from all the selected studies. The thescription of each column has been presented in chapter 4.3, Table 6.

Author(s) and Year	ASPECTS										
	Understanding of users	Learnability and Usability	Needed	Mutable	Appropriate	Aesthetic experience	Manageable	Privacy issues	Mean opinion score	Network - QoS	Others
Cheung et al., 2005		x		x			x				
Mahlke, 2002		x				x					
Muntean et al., 2006		x									
Knoche et al., 2007							x				
Colbert, 2004		x	x		x		x				
Bhattacharjee et al., 2003								x			
Wattimena et al., 2006										x	
Velsen et al., 2007						x	x				
Siller et al., 2003										x	
Ras et al., 2007		x	x		x						
Ali et al., 2005		x		x			x				
Wältermann et al., 2008									x	x	
Muntean et al., 2007									x	x	
Silzle, 2007		x			x		x				
Hamam et al., 2008		x			x	x	x			x	
Lu et al., 2008				x							
Zhang et al., 2007	x	x				x	x				
Collange et al., 2008										x	
Agrawal et al., 2007										x	
Monteiro et al., 2007									x		
Lopez et al., 2006									x		
Holone et al., 2008		x			x						
Vegiris et al., 2008	x		x				x				x
Ciubotaru et al., 2009						x					
Dyck et al., 2004										x	
Howes et al., 2007										x	
Siller et al., 2006										x	
Liu et al., 2007		x				x	x				
Calyam et al., 2008									x	x	
Zhang et al., 2009		x				x					
Garcia et al., 2009				x					x	x	
Moid et al., 2009										x	
Dufour et al., 2006										x	

Navarro-Ortiz et al., 2008		x				x			x		
Muntean, 2007		x		x			x				
Schaik et al., 2003		x					x				
Morris et al., 2001		x									
Trinh et al., 2009			x								
Cosma et al., 2008										x	
Yoo et al., 2008			x				x				x
Koivisto et al., 2007		x				x					
Pianesi et al., 2009		x				x					
Knoche et al., 2009							x				
Pandey et al., 2005		x									
Total: **44**											

Table 13: All aspects belonging to all studies.

Table 14 presents an overview of the purposes from all of the analysed studies. The description of each column has been prsented in chapter 4.4, Table 7.

Author(s) and Year	**Purpose**
Cheung et al., 2005	To examine the asymmetrical effects of negative and positive website attribute performance on satisfaction.
Mahlke, 2002	Examines the role of subjectively perceived factors of the experience of website usage in forming an intention to use a website.
Muntean et al., 2006	The aim of the experiment presented here was to investigate the usability and effectiveness of novel QoE layer enhancement for AHS(*), in the educational area. * Adaptive Hypermedia System (AHS)
Knoche et al., 2007	To determine the optimal zoom factors depending on the target device size in terms of its native resolutions. To give the participants a choice of following two clips playing in parallel by watching either the left or the right half of a screen.
Colbert, 2004	To compare students' experiences of communication before and during a rendezvous
Bhattacharjee et al., 2003	The focus was the analysis of models to enhance revenues from digital music sales in the presence of online music piracy.
Wattimena et al., 2006	To provide insight into the quality experience of gamers. A more specific goal is the development of an end-to-end quality measurement method that allows one to quantify the perceived quality of interactive gaming.
Velsen et al., 2007	Focuses on a specific form of secondary service, namely help facilities. Customers can visit a helpdesk to personally consult an agent about a problem or question they have. Or they can call a paid or toll-free telephone helpline for the same purpose.

Siller et al., 2003	It is believed that a better QoE can be achieved when the QoS and its interaction with the network and application layers is considered as a whole rather than as a single entity.
Ras et al., 2007	A capstone project helps the students to increase their problem solving competencies, improve their social skills (e.g., communication skills), and gather practical experience.
Ali et al., 2005	To show the benefits of a framework and the benefits of introducing a semantic service discovery and a trust approach for service selection. Focuses on the conceptual basis of their experiments rather than on understanding the low-level details of locating services.
Wältermann et al., 2008	This contribution introduces an ergonomic technique that aims at seamlessly switching the speech codec in Voice-over-IP calls during vertical handovers, based on SIP/SDP(*) session re-negotiation, the establishment of a parallel media stream and packet filtering. * The codec switching relies on the Session Initiation and Session Description Protocols (SIP/SDP).
Muntean et al., 2007	Focuses on the performance of multimedia streaming when using the Quality-Oriented Adaptive Scheme (QOAS) over an IEEE 802.11b Wireless LAN and compares it to that achieved when using other solutions that do not consider end-user quality in their delivery process such as TFRC(*) and non-adaptive schemes. * TCP-Friendly Rate Control Protocol (TFRC).
Silzle, 2007	The aim of the newly-generated taxonomy is to describe the components involved in the quality judgement process of Auditory Virtual Environments (AVE), and to quantify the relations between them for different applications.
Hamam et al., 2008	To propose a taxonomy for measuring the quality of experience of a haptic user Interface (HUI) applications and the focus on the haptics experience of the application apart from the video and audio feedback.
Lu et al., 2008	• Focus on the blocking of a single television channel for a particular user. • The authors investigate one important QoE measure, namely the content blocking probability.
Zhang et al., 2007	• Focuses on what are the most important usability factors for indicating the handset overall perceived usability. • How do the factors contribute to the handset overall perceived usability?
Collange et al., 2008	Focus on the correlations of QoE with end-to-end network level performance criteria and the authors propose a notion of sensitiveness to compare these correlations on different applications.
Agrawal et al., 2007	• Focuses on the emerging deployment of TV and video-on-demand services. • Focuses on a framework to aid planning and managing the deployment of IPTV services.
Monteiro et al., 2007	Proposes a subjective quality estimation tool for the continuous assessment of television over IP distribution systems. Methodology focuses the SVC encoder it can be extended to non-scalable video encoders.

Lopez et al., 2006	To define quality of metrics as well as methods to estimate them based on quantitative data observable by network operators. Focuses on the surroundings of the coding range CODVideo, that is a type-3 ISMA (*) profile. * Response Surface Analysis of type-3 Internet Streaming Alliance (ISMA) for multimedia delivery services over IP.
Holone et al., 2008	• Focuses on the algorithmic and technological aspects of the system. • Proposes a collaborative route planner. • Aims to find out how and why people would rate locations along the suggested routes. • Focuses on the following aspects of collaborative rating of accessibility: Efficiency, Effectiveness and Satisfaction
Vegiris et al., 2008	Focuses on validating an implementation of a state-of-the art audiovisual (AV) technologies setup for live broadcasting of cultural shows, via broadband internet.
Ciubotaru et al., 2009	• As QoE is difficult to assess, research has focused on proposing techniques to increase Quality of Service (QoS) level. • Proposes a Region of Interest Adaptive Scheme (ROIAS) for multimedia streaming, which differentiates the content adjustment process within the frame based on user interest in certain regions.
Dyck et al., 2004	To outline the basics of telepointers (*) and network issues, describing each of the HPT techniques in more detail, and then providing results from network simulations where the authors compare HPT to other common implementations. * Telepointers are valuable for supporting real-time collaboration.
Calyam et al., 2007	Focuses on only the most commonly used codecs i.e., G.722 voice codec and the H.263 video codec. At the same time it was focusing on the H.263 video codec at 768 Kbps dialing speed. * Recommendation G.722: 7kHz Audio-Coding within 64 kBit/s. * H.263 video codec standard.
Siller et al., 2006	Focuses on meeting and measuring the QoE of the end-user using an agent-based platform.
Liu et al., 2007	Focuses on the trustworthiness evaluation of service providers, since it applies in a similar way to service clients.
Calyam et al., 2008	Focuses on the KVMoIP(*) VNC(**) solution on both LAN and WLAN paths and focuses on the remote access of electron microscoper (i.e., remote microscopy). * Keyboard, Video and Mouse over IP (KVMoIP). ** Virtual Network Computing (VNC) solution.
Zhang et al., 2009	• Focuses intensively on how emotion is conveyed linguistically. • To support the real-time rendering of expressive characters and is the focus of the user study.
Garcia et al., 2009	Proposes a system that decides which type of wireless access network to connect with (for dual-band and tri-band devices), depending on the requirements of the IPTV clients, the available networks, and some network parameters, such as the number of loss packets and packet delay, to provide the maximum QoE to the customer.

Moid et al., 2009	The effectiveness of the proposed framework is demonstrated through the transmission of three test video sequences (Akiyo, Container, and Foreman), having different degrees of motion over an IEEE802.11 wireless network.
Dufour et al., 2006	To examine the behaviour of the Gnutella peer-to-peer file sharing network and propose a protocol modification to improve its performance. Furthermore, they propose to optimize Gnutella's performance by using a neighbour selection algorithm that considers topological information at the physical level as it forms the P2P overlay.
Navarro-Ortiz et al., 2008	• It aims at prioritizing real time services and at improving the performance of TCP application • This paper aims at fulfilling both objective (QoS) and subjective (QoE) requirements for conversational, streaming and interactive services, while best-effort fully utilizes the remaining capacity of the wireless medium.
Muntean, 2007	• To propose a performance-oriented Web user QoE model that apart from the user-related content adaptation, considers delivery performance-based content personalization in order to improve user experience when interacting with the system. • Proposed learner QoE model in adaptive and personalized education.
Schaik et al., 2003	A framework for the comprehensive investigation of response formats of online questionnaires is proposed as a basis for future research.
Morris et al., 2001	This research attempts to address gaps by introducing the construct of users' quality of experience as a potential mediator between the determinants of use and actual usage behaviours and their outcomes. Focus closely on perceived utility and the degree to which users are able to achieve intended outcomes from system usage.
Trinh et al., 2009	Propose equilibrium pricing strategies, taking into account the impacts of expected future incomes.
Cosma et al., 2008	The main focus was the estimation of the following metrics: One-way Delay (OWD), IP Delay Variation (jitter), and Packet Loss Ratio (PLR). With this range of parameters the authors can determine the ability, reliability and the robustness of each wireless technology.
Yoo et al., 2008	This study seeks to gain deeper understanding of how individuals experience time and space through the lens of tourism and this study seeks new insights as how to enhance tourists' experiences by supporting digital convergence and mobility.
Koivisto et al., 2007	The aim is to develop an integrated model for mobile information service adoption and use and apply it to internal communications of a community.
Pianesi et al., 2009	Focuses on issues relating to the core feature of the technology - namely, adaptivity - and to the interplay between the type scenario - that is, museum visits - and the role played by the guide.
Knoche et al., 2009	Focus on optimal trade-offs between size and resolution.
Pandey et al., 2005	Propose a new web crawler scheduling policy that prioritizes (re)downloading from the web, based on the expected gain in repository quality. They propose a novel approximation scheme for this purpose, coupled with an implementation

	technique in which measurements are taken in conjunction with index maintenance operations.

Table 14: An overview of the purposes of all the selected empirical studies.

Table 15 gives a short overview of the **results**, presented in each analysed study. The description of the columns (i.e. **author(s), year** and **results**) has been presented in chapter 4.5, Table 8.

Author(s) , Year	Results
Cheung et al., 2005	Provide partial support for the asymmetric effect on online user satisfaction. The positive performance of understandability, usefulness, and navigation had greater power than their negative performance, whilst the negative performance of reliability, access, and usability had a greater impact than their positive performance.
Mahlke, S. 2002	The four aspects of experience (i.e. usefulness (U), ease of use (EOU), hedonic quality (HQ) and visual attractiveness (VA)) succeed all in predicting the intention to use a website. The weight of contribution to the intention to use a website differs from U with the major influence to VA with the smallest effect.
Muntean et al., 2006	• Students from the QoE Adaptive Hypermedia Systems (QoEAHS) group had shorter *Study Session Time* than those that used the AHA! system. Therefore, the QoE-aware system has ensured a smooth learning process. • QoEAHA system provided better end-user satisfaction, than the AHA! system. It shows that the students considered the QoEAHA significantly more usable than AHA! • There is no strong correlation between learning performance and perception of usability. • An overall usability assessment shows that students considered QoEAHA significantly more usable than AHA!
Knoche et al., 2007	Display size represents a significant factor for the determination of optimal zooms. The zoom factors of 1.6 can have adverse effects on people's viewing experience of standard definition TV footage at QVGA size, the high end of the mobile TV resolution spectrum.
Colbert, 2004	• User experience of communication during a rendezvous does not appear to differ compared to before a rendezvous in terms of feeling of effort • User experience does not appear to differ in terms of satisfaction or convenience • The user experience of communication during a rendezvous was worse in terms of all other constructs i.e. social acceptability, frustration, disruption.
Bhattacharjee et al., 2003	Finds that revenue-maximization strategies for the seller do not necessarily involve efforts to eliminate online music piracy. Piracy reduction strategy is found to be different from revenue-maximizing strategy, especially when illegal online music is perceived to be of high quality.
Wattimena et al., 2006	Results demonstrate that ping and jitter have a significant negative effect on both the subjective (MOS) and objective gaming quality (kills), while packet loss goes unnoticed for values up to 40%. Especially, jitter in the network had a large negative effect on the perceived quality of the *Quake IV* players.

Velsen et al., 2007	It was a positive relation between the overall help facility quality and the customer's satisfaction with helpdesk. Helpline client's satisfaction will first and foremost be influenced by the quality of solutions offered. Organizations that want satisfied and motivated customers should pay attention and provide solutions of high quality of services such as helpdesk and helpline.
Siller et al., 2003	Results illustrate that by using network arbitration the Quality of Experience is improved. This implies that a better QoE can be obtained by the interaction of both Network and Application Layer arbitration. For example as better quality is provided for video, the QoE is improved.
Ras et al., 2007	Results shows that standard Wiki(*) functions improve communication and information sharing by means of explicit observation and experience documentation. The results implied that quality of experience was sufficient, but improvement is necessary. *An adapted free and open-source Wiki-based system called software organization platform (SOP).
Ali et al., 2005	Evaluating a transaction with a service is the cognitive process of making a judgment to assign a set of qualities for the experience that the user had with the service in that transaction. Furthermore, in experience, the qualities drawn from that experience are important. It shows that the second service has the highest trust value. This reflects the fact that service 2 costs less than the first service, and the user's desire for cheaper services.
Waltermann et al., 2008	Results showing that the proposed approach does not cause any interruption of the audio stream in about 90% of the test cases. PESQ(*) speech quality estimates reveal a quality advantage of 5.4% on average for the considered scenario. The procedure is constrained by the network handover latency, the network-inherent packet loss, and the jitter buffer status. The proposed Soft Codec Switching technique will positively contribute to mobile users'quality of experience and is therefore generally beneficial. *Perceptual evaluation of speech quality.
Muntean et al., 2007	The most significant results are obtained in highly loaded delivery conditions. For 15 clients when using Quality-Oriented Adaptive Scheme (QOAS), the end-user perceived quality is still at the 'good' quality level. Simulation results show that for the same average end-user quality, QOAS can accomodate a significantly higher number of simultaneous clients while also having higher total throughput. It is also significant to note that for the same number of simultaneous multimedia clients, the average values of the performance metrics such as end-user perceived quality, loss, delay and total throughput were always consistently better for QOAS than for other solutions. By using QOAS, an increased level of QoE can be provided to each viewer.
Silzle, 2007	Unknown ([*])
Hamam et al., 2008	Haptic User Interface (HUI) application was used to find criteria such as ease of use, modality choice, side effects/fatigue, user satisfaction, rendering quality, device stability and transparency, error rate etc. Results show that HUI is able to measure the quality of experience of the users.

[*] I have included this study with missing section about study results. Expecting to get the study results, I have emailed the author. The author has answered that the study is a PhD study and the results of this study are not published. Therefore, the results have not been reported in this thesis.

Lu et al., 2008	When the amount of users increases, there will be a point at which the blocking in P2PTV will be less than in IPTV, unless the IPTV network is extended accordingly. These results cannot only be used to analyse the blocking in existing systems.
Zhang et al., 2007	A statistical method, called Cronbach's Alpha, was used. Aspects such as Satisfaction, Controllability, Effectiveness, Frustation, Customizability, Navigation Attractiveness, Helpfulness and Consistency has Cronbach's Alpha coefficient ranged in value from 0.60 to 0.84 and average level of 0.70. This value has indicated an acceptable level of internal consistency.
Collange et al., 2008	Observe a 'desequencement' due to the retransmission (by TCP) of a lost packet, if at least on packet has been lost in the last congestion window. Observed that the response time is significantly influenced by the retransmission delays of lost packets. The authors proved the correlations between some traffic features and the network performance, i.e. the higher the loss rate, then the shorter and smaller the TCP connections.
Agrawal et al., 2007	The framework brings together a collection of models that capture the aggregate macroscopic behaviour of representative aspects of an IPTV service that impacts on service performance and hence the quality of experience from the service, like the viewer profiles, channel zapping delay, data server blocking probabilities, quality of video, etc. Convergence of residential services (voice, telephony, and internet access) over IP opens up new opportunities that blur the distinction between communication services providers.
Monteiro et al., 2007	The subjective quantification of the impairments caused by changing the temporal resolution, the spatial resolution and the Root Mean Square of the Error (RMSE) in an encoded video. It is possible for an encoder (or video server) to select the best cost-benefit ratio, i.e. try to maximize the QoE according to the bandwidth available or the minimum bit rate for a specified level of quality. Shows that an increase in the Y-PSNR(*) of an encoded video does not necessarily correspond to an increase in the subjective quality of a viewer. Therefore, the estimated QoE must be considered as an important metric for the design and assessment of the encoder quality layers. *Peak Signal-to-Noise Ratio (PSNR).
Lopez et al., 2006	The behaviour of multimedia systems delivered over IP networks. • The most prominent sub-aspects affecting the observed response are the delay and the video rate. However, all the considered factors are statistically significant. • ANOVA analysis reveals that both CODVideo and Delay play an important role. It has to be pointed out that, according to the R-squared value, the model fits quite approximately the real behaviour of the response variable QoE_{Video}. • The results reveal that H.264/AVC(*) transported over RTP_2 offers a significant improvement in the exhibited resilience against packet losses, being able to tolerate ut to 0.1% packet losses. The factors CODVideo and Delay, however do affect the final video quality QoE_{Video}. * ISO /IEC14496-10 (video codec recommendation).
Holone et al., 2008	The users were interviewed about the route they had followed: 'Route quality' – the quality of the route is regarded as good. The route planner 'knows best', and that they have been guided along the shortest path to solve the task

	Usefulness of the service – The trend of the usefulness ratings is positive. Some users rated the usefulness of the system high. User satisfaction – Satisfaction from the perspective of the tools perceived usefulness should intuitively increase as the number of non-negotiable obstacles decrease. Interestingly, even when participants rate the usefulness lower, they tend to think about it as a useful service for others.
Vegiris et al., 2008	Subject evaluated the quality of video as 'very good' or 'excellent', more than 80% and the quality of sound more than 70%. More than 60%, evaluated the degradations and impairments of the video and audio as 'perceptible, but not annoying', or 'imperceptible'. Lastly, the total experience of the event was evaluated as 'very realistic' or 'interesting' by more than 70%. Evolution of the present model could include the incorporation of such subjective tests and perceptually adapted metrics into the performance definition QoS/QoE of a system, the extension of the properties for the subsystem parameterization as well as audio performance estimation.
Ciubotaru et al., 2009	The results are presented in terms of two objective metrics which estimate user perceived video quality: Pick Signal-to-Noise Ration (PSNR) and Video Quality Measurement (VQM) show how by using Region Of Interest-based Adaptive Scheme (RIOAS) there is a clear benefit in terms of quality in the areas of highest user interest. Results show that how users benefit by using Linear-ROIAS and perceive higher quality in the area they have the highest interest in.
Dyck et al., 2004	• The latency of UDP-based techniques is not affected significantly by loss, but rather depends on available bandwidth. As bandwidth decreases, latency remains low. • When not predicting, error arises from jumps in the telepointer path due to lower send rates, jitter and loss. • The error from prediction was more significant at higher loss rates since more prediction is occurring during burst loss. • The effect of measurement noise on accuracy is substantial.
Calyam et al., 2007	The closed-network methodology (*) leveraged test case reduction strategies that significantly reduces a human subject's test duration without compromising the rankings data required for adequate model coverage. Additional variants need to be derived for accurately estimating end-user VVoIP QoE because the network performance bottlenecks manifest differently at higher dialing speeds and are handled differently by other video codecs. * The closed-network methodology focused on the H.263 video codec at 786Kbps dialing speed.
Siller et al., 2006	• User's expectations were met (good), indicating a general level of satisfaction rather than disapproval. • Answer shows that user's understanding of QoE was 'Something related to media quality provided as a result of the request'. • The platform objective is to meet the QoE requirements during the transmission and it generates a mapping of the QoS metrics to QoE. It employs both objective and subjective weighting factors. • QoE requirements can differ from one user to the other. It is believe that by empowering the user to select the quality according to expectation, QoE can be improved relative to the individual.
Liu et al., 2007	A distributed reputation mechanism was presented for evaluating the trustworthiness of

	a service provider in ubiquitous computing environments. The proposed reputation mechanism can easily apply to trustworthiness evaluation of service clients. Quality of Experience evaluation depends on other factor such as timeliness of payment.
Calyam et al., 2008	The impact of network connection quality on the user QoE MOS(*) in a remote microscopy session. It shows that novice rankings are relatively more liberal than expert rankings due to the inherent intensity of the actions involved. Difficulty level of expert user tasks is higher than novice user tasks, because MOS ranking dip is higher in the expert user cases than the novice user cases. Throughput trends are significantly less dense in case of the Direct GigE connection as compared to the Public LAN connection. Inverse relationship between the network connection quality and user control effort is a driver, where a user expends more effort (i.e., mouse moves/clics and keyboard strokes) on poor network connections. The real-time control and video image transfer traffic is extremely bandwidth intensive for achieving 'at-the-microscope' QoE. *Mean opinion score.
Zhang et al., 2009	Consistently between the groups the quality of social interaction improved with the 3D condition. E-drama provides an opportunity for the developers to explore how emotional issues embedded in the scenarios, characters, and dialogue can be represented visually. E-drama system creates a new channel for young people's classroom communication. E-drama shows that application of expressive characters to online role-play contributes positively to an already engaging user experience.
Garcia et al., 2009	Three wireless technologies: IEEE 802.11a, IEEE 802.11g and WiMAX. Jitter testing – the average jitter was very similar in WiMAX and IEEE 802.11a. These technologies are approximately 35% better than IEEE 802.11g. The reason for that is IEEE 802.11g seems to be less stable. Lost packet testing – Results shows that IEEE 802.11a was less robust than other ones. Effective bandwidth – The theoretical number from test shows that IPTV content distribution system can support 5 IPTV channels in the WiMAX scenario. QoE comparison – The IEEE 802.11 technology has a very stable network's QoE.
Moid et al., 2009	Tested - The three video sequences: Akiyo, Container and Foreman. Results indicate that adapting the transcoding parameters at the application layer to the current channel condition does not produce significant processing delays. Objective and Subjective Video Quality – the data-link layer optimization exhibits higher packet processing time. The foregoing confirms that the proposed framework demonstrates a tradeoff between the PSNR (*) gain and the increase in packet processing time. The proposed framework will support efficient streaming of video over IEEE 802.11 wireless networks. *Peak-signal-to-noise-ration (PSNR) is the metric to measure the quality of reconstruction in image compression.
Dufour et al., 2006	Simulation results show that the proposed neighbour selection algorithm to form the Gnutella (*) overlay based on physical topology information achieves a greater QoE for users by reaching more nodes faster with queries and receiving query hits more quickly. *Gnutella peer-to-peer file sharing network.

Navarro-Ortiz et al., 2008	A solution for the IEEE 802.11e HCCA (*) mechanism, which is designed for a heterogeneous network. Results show that VoIP maintains good quality (MOS is approximately 4.5) for the different schedulers up to the maximum number of admitted VoIP users. The Monolithic Shaper-Scheduler (MSS) with territory method presents: • The quality perceived by a conversational user is less influenced by the traffic generated by other services. This has been illustrated for both objective (maximum delay) and subjective (MOS) metrics. • TCP applications are handled efficiently. Interactive services will take the necessary resources to achieve the require performance, and best-effort will fully utilize the unused capacity. • End-to-end impairments (i.e. delay, bottlenecks, packet losses) are considered so that the IEEE 802.11e wireless network reacts against these impairments, maintaining the demanded quality. Finally, the authors' solution was compared to other HCCA scheduling algorithms, the reference scheduler and Fair Hybrid Coordination Function (FHCF), using VoIP and FTP traffic. The results show that the combination of the MSS and the territory method obtains higher system capacity for VoIP traffic (up to 32 users). *Hybrid coordination function Controlled Channel Access (HCCA).
Muntean, 2007	Learning achievement – Results show that the students from QoEAHA group allocated a lower period of time than AHA! group, and students from QoEAHA succeded to focus better on the studied material. It means QoEAHA system provided a better QoE for the learners and students improve their satisfaction. Aspects such as navigation and presentation achieved similar results for both groups. This shows that the QoE enhancements did not affect them. Learning performance for QoEAHA group has significant better end-user QoE than group AHA!
Schaik et al., 2003	Factorial analyses have shown three scales such as disorientation, perceived ease of use and intensity of flow and these scales had high reliability. Validity was high on these scales.
Morris et al., 2001	A factor structure matrix for the 'high quality of experience' and 'low quality of experience' items was used. Findings show that items were more important to perception of high quality of experience than another one. It shows that ease of use and usefulness are significant determinants of IT use and there is providing for face validity of the quality of experience construct.
Trinh et al., 2009	Socio-economic aspects for future communication networks such as cost, pricing model of providers are important in the marketplace. There are also differences between short-term and long-term in a dynamic market. For example, in the short term, profit is more significant, which means that discount factor is high and the incumbent ISP tries to hold lowering prices.
Cosma et al., 2008	Analysis was based on statistical interpretations of three following metrics: One-Way-Delay (OWD), Packet Loss Ratio (PLR) and Inter-Packet-Delay Variation (IPDV). Values such as low delay and jitter of the ethernet client were delivering a high QoE for audio and video transmissions. The client Wifi 802.11 performance was very good and it had 0% packet loss for the audio stream and 0.034% for the video stream. Using this technology, QoE was excellent on audio and video transmissions.

	The performance of the WiMax client also offered satisfactory results for audio, but on the video channels was some interruptions and image blockiness. Therefore, reliability was lacked for the WiMAX client. The UMTS client has bandwidth limitations, therefore, QoE was satisfactory with the lower quality settings.
Yoo et al., 2008	This study provides a scenario for future mobile 2.0 service developments. The authors maintain that they are still very early in their analysis. They believe that three elements of spatiotemporal narrative can be a useful conceptual device to understand how individuals experience time and space, meeting different fundamental human needs. Second, this analysis suggest a conceptual nature of mobility and our experience of space, through the lens of spatiotemporal narratives.
Koivisto, 2007	Analysis was based on acceptance of mobile phone mediated communication technologies in community context. Findings show that three sub-communities such as Spanish, Finnish and Chinese students, have similar acceptance patterns. 'Usefulness' was more important than the 'ease of use' in the acceptance of basic communication services. There was also a strong correlation between attitude and intention. The factor 'time' was not important in relation to the learning curve, but it was important in system development.
Pianesi et al., 2009	Results show that motivation to use museum guides is mainly utilitarian. Adaptivity issues can summate that the perceived control on the interaction plays insignificant role for the acceptability of adaptive museum guides. Personalization impacts on intrinsic, but not extrinsic, motivations. The impact of the control issue on the acceptability of the adaptive guide is both low and indirect. It primarily mediates the effects of the global experience on the perceived ease of use. The driving force for acceptance being the guide's Perseived Usefulness in improving the interaction with the museum.
Knoche et al., 2009	Results show that actor size on the minimal angular resolution was insignificant effect. Low-resolution mobile content the study shows a large effect of the preferred viewing size. Multimedia models for video content on mobile devices might be displayed an adequate size, because right size is significant. The acceptability gains eXtreme Long Shots (XLS) by zooming were substantial for viewing ratios larger than 8.5. Lab experiments may be a conservative estimate of the acceptability of video quality as consumed by people on the move (e.g. on a train).
Pandey et al., 2005	The authors devised an efficient method and together with user-centric scheme attained over 99% quality on the user experience. They have been introduced a new web crawling paradigm, which were designed for search engines. Intention for that was to allocate resources to crawling tasks, which will be possible to maximize the quality of the user experience. Findings show also that this scheme requires fewer resources to achieve the same user experience quality.

Table 15: An overview of results from all selected empirical studies.

Appendix G – A short synopsis of all analysed studies.

Table 16 presents an overview of all analysed studies.

- The column **paper** refers to the B list of references.
- The column **year** shows the year of publication.
- The column **subject** is splited into two columns **with** and **without**. These columns present what kind of subjects participated in these studies (i.e. with human participants and without human participants.
- The column **type of study** presents the type of each study (e.g experimental, case study etc.).
- The column **aspects** presents how many aspects have been analysed by the authors per each selected study.
- The column **purpose** presents the purpose of ech study.
- The column **results** presents the sub-aspects from each study..

Note. It is important to get ecquained with Table 17 before reading the Table 16.

Paper	Year	Subject		Type of study	Aspects	Purpose	Results
		With	Without				
[1]	2005	■	□	Survey	L M F	6	R - 1 R - 3 R-X
[2]	2002	■	□	Survey	L E	6	R - 3 R - 4 R-X
[3]	2006	■	□	Exp.	L	1	R - 1 R - X
[4]	2007	■	□	Exp.	F	7	R - X
[5]	2004	■	□	Exp.	L N A F	1	R - 1 R - 8 R-X
[6]	2003	■	□	Exp.	P	4	R - X
[7]	2006	■	□	Exp.	Q	3	R - 2 R - 5 R-X
[8]	2007	■	□	Survey	E F Q	1	R - 1 R - X
[9]	2003	□	■	Exp.	Q	2	R - X
[10]	2007	■	□	Case	L N A	1	R - 8 R - X
[11]	2005	□	■	Exp.	L M F	4	R - X
[12]	2008	□	■	Exp.	S Q	2	R - 2 R - 5 R-X
[13]	2007	□	■	Exp.	S Q	2	R - 2 R - 6 R-X
[14]	2007	■	□	Survey	L A F	3	Unknown
[15]	2008	■	□	Case	L A E F Q	2	R - 1 R - 4 R-X
[16]	2008	■	□	Case	M	4	R - X
[17]	2007	■	□	Survey	U L E F	9	R - 1 R - X
[18]	2008	□	■	Exp.	Q	1	R - 2 R - X
[19]	2007	■	□	Case	Q	1	R - 8 R - 9 R-X
[20]	2007	■	□	Survey	S	2	R - 7 R - X

[21]	2006	□ ■	Case	S	2	R - 2 R - 6 R-X
[22]	2008	■ □	Exp.	L A	3	R - 1 R - 3 R-X
[23]	2008	■ □	Exp.	U N F O	3	R - 9 R - X
[24]	2009	□ ■	Exp.	E N A F	2	R - 9 R - X
[25]	2004	□ ■	Exp.	Q	3	R - 2 R - 5 R-7 R-X
[26]	2007	■ □	Exp.	Q	2	R - X
[27]	2006	■ □	Survey	Q	1	R - 1 R - X
[28]	2007	□ ■	Exp.	L E F	4	R - X
[29]	2008	■ □	Exp.	S Q	3	R - 7 R - X
[30]	2009	■ □	Exp.	L E	3	R - 8 R - X
[31]	2009	□ ■	Exp.	M S Q	8	R - 2 R - 5 R-7
[32]	2009	□ ■	Exp.	Q	2	R - 2 R - 6 R-X
[33]	2006	□ ■	Exp.	Q	4	R - X
[34]	2008	□ ■	Exp.	L E S	3	R - 2 R - 6 R-X
[35]	2007	■ □	Exp.	L M F	1	R - 1 R - X
[36]	2003	■ □	Survey	L F	1	R - 4 R - X
[37]	2001	■ □	Survey	L	5	R - 3 R - 4
[38]	2009	■ □	Case	N	5	R - X
[39]	2008	□ ■	Exp.	Q	8	R - 1 R - 2 R-5 R-6 R-7 R-X
[40]	2008	■ □	Case	N F O	3	R - X
[41]	2007	■ □	Exp.	L E	1	R - 3 R - 4 R-X
[42]	2009	■ □	Exp.	L E	10	R - 3 R - 4 R-X
[43]	2009	■ □	Exp.	F	7	R - 9 R - 9
[44]	2005	□ ■	Exp.	L	5	R - X

Table 16: An overview of all selected studies

Table 17 presents all symbols and descriptions, which are used in Table 16.

- The column **symbol** presents acronyms, which are used in Table 16.
- The column **description** gives more explanation about each symbol.

Symbol	Description
[1]	The number with bracket means selected and analyzed study. Total number was 44 (see References B)
2008	Year
	Subject
■	Market square
□	Unmarket square
	Type of study
Exp.	Experiment study
Survey	Survey study
Case	Case study

	Aspects
A	Appropriate
L	Learnability and usability
N	Needed
M	Mutable
E	Aesthetic experience
F	Manageable
P	Privacy issues
S	Mean Opinion Score
Q	Network - QoS
O	Others
U	Understanding of users
	Purpose
1	Education and communication
2	Multimedia streaming
3	Real-time collaboration
4	Resource sharing and trustworthiness
5	Future communication network
6	Web application
7	Size and resolution
8	Wireless Network
9	Relationship between factors and design
10	Future of technology
	Results (sub-aspects)
R - 1	Satisfaction
R - 2	Loss/packet loss
R - 3	Usefulness
R - 4	Ease of use
R - 5	Jitter
R - 6	Delay
R - 7	Bandwidth
R - 8	Communication
R - 9	Quality of video
R-X	Others

Table 17: Descriptions of all symbols