On the notion of quality of experience

Kalevi Kilkki
Aalto University, Espoo, Finland

kalevi.kilkki@aalto.fi
kalevi.kilkki@gmail.com

Abstract

This report provides a comprehensive review of research activities using the concept of quality of experience. Before 1990, the quality of experience was primarily used in philosophical and psychological articles, where it appeared occasionally. Since the 1990s, the term has mainly been used in business analysis, first in the tourism sector, then in human-computer interaction studies, and finally in the context of communication networks. In the networking field, Quality of Experience (QoE) usually refers to user satisfaction measured on the mean opinion score (MOS) scale. Other fields of study use the concept more widely covering all kinds of experiences and their consequences. Therefore, more active cooperation between the fields would be beneficial. To this end, this report introduces a framework covering all fields and uses of quality of experience from networking technology and human-computer interaction to business development and the well-being of society. Finally, seven recommendations are provided to facilitate the quality of experience research in the future.

Keywords: quality of experience · quality of service · satisfaction · mean opinion score · utility

Contents

Preface

This report is a summary of fifteen years of considerations related to the concept of quality of experience. I gave the first public presentation in which I discussed quality of experience in June 2007.[1] By that time, I had already worked for about 20 years to design, model, and optimize traffic control mechanisms in communications systems. The adopted viewpoint and the main results were technical with an economic flavor.[2] I had become somewhat dissatisfied with the limited viewpoint and methods used in technical research and development. Thus, when quality of experience emerged as an alternative to the more technical term quality of service, I was keen to express my opinion about the relationship between quality of service and quality of experience. In addition to the presentation, I wrote an article published in 2008 [173]. Over the last 14 years, I have closely followed quality of experience research and participated in inspiring Dagstuhl seminars [72, 91] but I have not participated in any research project with quality of experience as a key theme. Moreover, as a university lecturer, I have spent most of my working time teaching, instructing students, and writing course material.

Four years ago, I decided to write a kind of summary of the three main research topics in my scientific career, quality of service, value of time, and quality of experience. Quality of service and value of time are dealt with in two papers, [174] and [175], respectively. With quality of experience, I had a more ambitious objective, a high-quality journal article, but that goal was not realized in the expected format. The first version of the manuscript was bluntly rejected, and the next two versions were heavily criticized; it seemed impossible for me to get the manuscript accepted before my scheduled pension date. Hence, I decided to write a report rather than a peer-reviewed article. Here is the result.

The objective of this report is to express my best understanding of the quality of experience from several perspectives. The body text consists of three parts: history, philosophy (or framework), and practice (or mean opinion score). The historical part is based on an extensive collection of papers and the use of simple statistical analysis related to the number of authors, the use of different research methods, the citations to key authors, etc.[3] The philosophical part aims at providing a balanced account of the use of quality of experience in different fields of study.[4] The practical part discusses the problems of the mean opinion score when it is used as a measurement scale for assessing the quality of experience.[5] Moreover, a brief section presents seven recommendations for the next generation of quality of experience researchers.

I have also included some additional material collected over the past 12 years in seven appendixes.[6] Appendix 1 offers a small survey of the use of quality of experience in literature. I would encourage all readers interested in quality of experience to read through the 36 excerpts from 18 books, because they demonstrate the meaning of the phrase "quality of experience" when it is not used as a specific scientific or technical term. Appendix 1 also presents a list of 60 quality of experience articles. The articles are

[1] EuroFGI IA.7.6 Workshop on Socio-Economic Issues of NGI, Santander, Spain, 28-29 June 2007.

[2] Results include a doctoral thesis (Traffic characterisation and connection admission control in ATM networks), patents (e.g., US patent 6047326, Accounting system and method for a nominal bit rate network service), and a book (Differentiated Services for the Internet).

[3] The historical part of this report was already included in the paper that went through two rounds of a review process by a recognized journal. The reviewers' comments have been addressed in this report. I would like to thank the reviewers for their valuable remarks.

[4] Maybe half of the philosophical part has been written after the last review process. Thus, every reader should be cautious with this part – of course, I have done my best to write as consistent and truthful text as possible.

[5] The MOS section should be considered new material without any peer review, although some parts of the text have gone through a review cycle.

[6] As to the appendixes, there has not been any peer-review whatsoever. Thus, enjoy but be cautious.

divided into 6 groups of 10 articles in a way that the first group, in my mind, consists of the most recommendable articles, the second group the next 10 most recommendable articles, etc. Appendix 2 provides an estimate of the total number of the quality of experience papers with at least one citation in Google Scholar. Appendix 3 compares the vocabularies used in different fields of study. Appendix 4 provides statistics on the use of different objective methods to assess the quality of experience. Appendix 5 shows the results of two small experiments related to the mean opinion score carried out at Aalto University. Appendix 6 shows the results of DRM (Day Reconstruction Method) studies conducted with students at Aalto University. Finally, Appendix 7 provides some considerations about the meaning of concepts like utility, usefulness, value, and benefit.

If you find this report interesting and valuable, please feel free to disseminate it to anyone interested in the concept of quality of experience. This report can be used and distributed under the license Attribution-ShareAlike 4.0 International (CC BY-SA 4.0).[7]

As for the possible continuation of this endeavor to dissect the concept of quality experience, I cannot promise much. It is always difficult to predict whether there is enough motivation to return to a finished document, but now the task is even more difficult because I am entering a new phase in my life. In any case, I welcome all forms of feedback.

Helsinki, June 2022

Kalevi Kilkki

[7] https://creativecommons.org/licenses/by-sa/4.0/

1 History

This section describes how the use of quality of experience as a scientific concept has evolved. The discussion covers various fields of science and technology from philosophy to network engineering.

The renowned philosopher John Dewey wrote in 1887 [76]: "The experience, as an existence at a given time, has for ever vanished. Its meaning, as an ideal quality, remains as long as the mind does. Indeed, its remaining *is* the remaining of the mind; the conservation of the ideal quality of experience is what makes the mind a permanence."[8] One hundred and thirty-one years later, Redowan Mahmud with four co-authors wrote [193]:[9] "In brief, MCC combines cloud computing, mobile computing and wireless communication to enhance Quality of Experience (QoE) of mobile users and creates new business opportunities for both network operators and cloud service providers."

John Dewey was a pragmatist philosopher whereas Mahmud and his colleagues have a degree in computer science. The unifying factor between the papers is quality of experience. The following story-line between the two papers consists of research carried out in several disciplines including philosophy, psychology, tourism, recreation, user experience in the context of human-computer interaction (HCI), and quality of service (QoS) in communication networks.

Material and research method

The main research method used in this report has been to collect a large set of papers with quality of experience in the text (excluding references and biographies).[10] A paper refers to a journal article, a paper published in a conference publication, or a book chapter. In contrast, academic dissertations, textbooks, and patents have been omitted from the collection, although some of them are used as references. Furthermore, two specific cases containing the phrase quality of experience have been omitted: (the quality of) experience products and (the quality of professional) work experience. In both cases, the phrase "quality of experience" has been as used in a different sense than what is the focus of this report.

Since the number of scientific papers mentioning quality of experience is huge, additional limits were necessary. If a concept appears only once or twice in a paper, it is likely a minor issue in the paper. In addition, if a paper has been cited only a few times, the influence of the paper on other researchers has likely been weak. Therefore, the focus of this study has been on finding papers that contain at least three sentences with quality of experience or QoE. Another criterion has been to limit the search to papers with at least 50 citations.[11]

Consequently, the numerical analysis in this report is based on three sets of papers:

Set Q1: The quality of experience is mentioned at least once without any citation criterion,

Set Q3: The quality of experience is mentioned at least three times without any citation criterion, and

Set Q3C: The quality of experience is mentioned at least three times with the following citation criterion: At least 50 citations if the paper was published in 2017 or before, 40 citations for papers published in 2018, 18 citations for papers published in 2019, and 10 citations for papers published in 2020.[12]

[8] Italics in the original.
[9] This paper was selected as an example because it is the most cited paper published in 2017 or later that uses quality of experience as an essential concept. The main topic of the paper is fog computing.
[10] However, if quality of experience appears only in references or biographies, the paper is not included in the set of quality of experience papers.
[11] Citation data refer to the number of citations given by Google Scholar, https://scholar.google.com/. However, other sources have also been used for collecting relevant papers.
[12] The logic of the limits is that the number of included papers is about 50 per year. Most of the citation data used in this report was collected on 20 November 2021.

Note that all papers in Set Q3C belong to Set Q3 and all papers in Set Q3 belong to Set Q1. The basic data on the articles in the three sets are presented in Table 1.

Table 1: Basic information about the papers in Sets Q1, Q3, and Q3C. HCI = human-computer interaction. Note: the row "Quality of Experience" shows the data of those papers that use capitalization instead of form quality of experience.

				Set Q3C				
	Set Q1	Set Q3	Set Q3C	Philos.	Psych. & education	Tourism & recreation	HCI	Networking
Number of papers	1314	1026	689	11	73	34	73	480
referred in this report	208	181	140	3	15	11	29	79
recommended*	60	57	55	3	8	7	16	25
Average number of								
authors	3.58	3.87	4.03	1.45	2.23	2.38	4.63	4.43
citations	166	129	182	567	384	555	169	120
First author								
female	22%	21%	20%	-	37%	53%	24%	14%
from a private firm	8%	8%	7%	-	-	-	5%	10%
Papers								
oldest publishing year	1887	1948	1964	1971	1989	1995	1996	2001
aver. publishing year	2010	2012	2012	2003	2004	2006	2013	2014
with formulas	39%	44%	46%	27%	1%	-	38%	58%
with QoE	67%	81%	78%	-	3%	15%	78%	98%
Quality of Experience	36%	43%	38%	-	4%	9%	37%	47%
Affiliation (first author)								
Europe	44%	46%	43%	9%	42%	29%	52%	44%
USA & Canada	30%	25%	26%	91%	53%	41%	30%	19%
China & Hong Kong	12%	14%	17%	-	-	9%	5%	22%
Other countries	14%	15%	14%	-	4%	21%	12%	15%

* The list of 60 recommended papers is presented in Appendix 1. All 60 papers are included in the last five columns (one paper, [236], is classified as other business).

The number of citations varies enormously. The most cited paper in Set Q1 is Ryan & Deci [254], which has gathered over 22000 citations. In Set Q3C, the most cited paper is a tourism-related paper with 5100 citations [24]. In the case of the technical fields (HCI and networking), the most cited paper [209] is an image-quality paper with 2740 citations.

As shown in Table 1, there are substantial differences between the fields in the use of capitalization (Quality of Experience instead of quality of experience) and the use of the acronym QoE. Of the 480 networking papers in Set Q3C, only ten papers do not use the acronym QoE. On the contrary, among the 73 HCI papers in Set Q3C, 27 papers use the format Quality of Experience with QoE, 30 papers use the format quality of experience with QoE, and 16 papers use the format quality of experience without QoE. In all other fields combined, only 12 papers (out of 136) use the acronym QoE.

The main reason for the occasional use of QoE in tourism papers seems to be the prominent paper by Otto & Ritchie [228]. Their paper was the first to use the acronym QoE (in the form of QOE).[13] However, the most common form in their paper was the quality of service experience while the acronyms were used in a table comparing QOS and QOE frameworks.

In psychology, the acronym QoE is not used except in some education-related papers that typically fall somewhere between HCI, education, and psychology. For instance, [110, 189, 314] have studied experiences with e-learning systems. As to the origin of QoE in these papers, Hameed et al. [110] did not elaborate on the background of QoE, Ljubojevic et al. [189] referred to Alben [9], and Vasileva-Stojanovska et al. [314] referred to ITU-T[14] standard P.10/G.100 [143]. Thus, it seems that the occasional use of the acronym QoE in psychology and education originates from networking research. Although the eminent psychologist Mihaly Csikszentmihalyi regularly used the term quality of experience, it seems that he never used the acronym QoE. In networking, the main reasons for the prevalent use of QoE are the popularity of acronyms in general [88] and the widespread use of the acronym QoS before the emergence of QoE [174].

Philosophy

The oldest paper in Set Q3 is a rarely cited paper published in 1948 deliberating the essence of experience [89]. In Set Q3C, the oldest paper was published in 1964 [149]. Both papers are based on John Dewey's insight into experiences. The first two papers in which quality of experience appears in the title were Stephen Crites's *The narrative quality of experience* [61] and Gilbert Harman's *The intrinsic quality of experience* [112].[15] However, two issues diminish the importance of these remarkable papers in the context of this report. First, no paper in Set Q3 refers to Crites, and only one paper refers to Harman. Second, quality of experience is not a specific term in these papers. Instead, the format is 'the narrative quality' of experience and 'the intrinsic quality' of experience. These two papers did not have any significant effect on the later usage of quality of experience.

The situation has remained unchanged in the domain of philosophy over the last fifty years: although some philosophers occasionally mention the quality of experience, it has not become an established philosophical concept. The only significant exception is Giulio Tononi, who parallels quantity of experience and quality of experience when discussing the fundamental nature of consciousness [304]. Tononi's integrated information theory framework even allows mathematical treatment of both concepts. Nevertheless, Tononi's ideas have not had any significant effect on the treatment of quality of experience in other research fields.

Psychology

Optimal experience in work and leisure by Csikszentmihalyi & LeFevre [66], published in 1989, was the first paper in which the quality of experience was clearly specified and that was widely cited in the subsequent literature. Their main research tool is the Experience Sampling Method (ESM). The method is based on randomly allocated self-reports describing a person's momentary situation including the type of activity, location, companionship, and the quality of experience. Within their framework, the quality of experience is a construct that consists of two main dimensions, affect and potency, and two less homogeneous dimensions, cognitive efficiency and motivation. *Affect* defines how happy or satisfied the person is, *potency* defines the level of arousal or excitement, *cognitive efficiency* measures the level of concentration, and *motivation* measures whether the person wishes to do what one does.

In this kind of approach, the quality of experience is a construct built on questionnaire answers. The main open issues are how the questions are composed, how they are grouped, and how the groups are labeled. When these issues are resolved and one group is named quality of experience, the concept is

[13] More accurately, [228] is the oldest paper in Set Q1 using the acronym QoE or QOE; I am unaware of any notable, older paper using abbreviation QoE.
[14] ITU-T: International Telecommunication Union, Telecommunication Standardization Sector.
[15] Harman's paper [112] is still the most cited paper in which quality of experience appears in the title.

unambiguously defined and operationalized. This apparent clarity is helpful with regard to scientific research and writing academic papers. On the other hand, there is no compelling reason to contemplate the fundamental nature of experience. Nonetheless, some psychology papers offer interesting deliberations about experiences. For instance, Csikszentmihalyi & Schiefele [67] cite Dewey, Joyce, Jung, and Wittgenstein when discussing the nature of experience in the art education. Melzack [204] discusses the nature of experience in the context of pain and reminds us of the fact that the brain can generate any kind of quality of experience without any sensory input.[16]

The most important applications of Csikszentmihalyi's psychological treatment of experience are education [205, 263, 273] and flow [19, 178, 223]. The impact of the results provided by Csikszentmihalyi is illustrated by the fact that 65 papers in Set Q3C refer to him, including papers about tourism [60, 228], HCI [95, 256], and networking [73, 173].

In a technical context, research focus is often to avoid technical deficiencies that result in user dissatisfaction. In contrast, the domain of positive technology attempts to develop means to improve human well-being. Riva et al. [252, figure 1] divide the efforts of positive technology into three areas: emotional quality (arousal, valence, object), engagement/actualization (challenge/skills, goals, presence), and connectedness (collective intentions, social presence, empathy). All of these aspects are essential when the research objective is to enhance experiences rather than to avoid dissatisfaction.

Tourism and recreation

Tourism and recreation combine strong business drives with emphasis on rich experiences. Quality of experience, therefore, seems to be a plausible term when analyzing customer behavior. Nonetheless, it is worth noting that quality of experience has various connotations particularly for those that have not used it in a formal context. For instance, Arimond argues in his doctoral thesis [16] that "In addition, managers began using the jargon phrase "quality of experience." However, this approach [measuring hunter satisfaction] did not really identify a method for assessing the characteristics of a quality hunting experience."

When a term becomes common in scientific or business literature, it tends to crystallize to some specific meaning. Arimond's thesis was published in 1986, whereas the oldest tourism paper in Set Q3C, Crompton & Love [62], was published in 1995. Crompton & Lover referred to Brown [46] as the source of the concept, even though Brown mentions it only in two sentences. Other important early quality of experience papers in the tourist sector were Baloglu & McCleary [24], Otto & Ritchie [228], and Tian-Cole et al. [302]. These papers formed the basis for further studies in analyzing tourist and recreation experiences. Several early papers also referred to classical service quality papers, such as the prominent paper by Parasuraman et al. [230].

Baloglu & McCleary [24] define the *quality of experience* as the label for the first and most important personal/cognitive items while the two other items in the same category, *attractions* and *value/environment*, are less important. Their model contains two other endogenous variables, *affect* and *overall image*. Regardless of the psychological undertone of quality of experience, the factors contain material items like hygiene and cleanliness, quality of infrastructure, and suitable accommodations. These specific characteristics of their model limit its direct applicability in other fields of study.

Tian-Cole et al. [302] define four variables for modeling visitors to a wildlife refuge: quality of performance, quality of experience, overall satisfaction, and behavior intention. The four indicators of quality of experience are achievement, introspection/nostalgia, escape, physical fitness, and new people. An interesting finding is that visitor satisfaction and (the assessment of) service quality are independent of each other. Moreover, they conclude that the overall service quality and overall satisfaction are not the same construct. This is understandable in the sense that the overall satisfaction depends more on

[16] Mlodinow states [211, p. 45]: "But modern neuroscience teaches us that, in a way, all our perceptions must be considered illusions:" Every conscious perception requires complex processing in the brain that can be called a simulation. The simulation creates all perceptions with or without any stimulus from reality - whatever reality means.

nature itself than on the service provided by the refuge management. Certainly, bad service can destroy all experiences, but real satisfaction is usually obtained from the genuine content be it a wildlife refuge or a video call.

After these pioneering papers, quality of experience has often been used as a name for a variable in statistical analysis. For instance, Kivela & Crotts [177] define a dependent variable by the question: "Overall, what contribution did food and cuisine have on your overall quality of experience as a tourist visiting Hong Kong?" Xu & Chan [338] discuss hotel experience and divide brand knowledge into brand awareness, brand associations, and quality of experience.

In the analysis of tourism business, experience quality is often used instead of quality of experience without making any difference between them. For instance, Chang & Horng [52] define experience quality as "representing how customers emotionally evaluate their experiences as they participate in consumption activities and interact with the service surroundings, service providers, other customers, customers' companions, and other elements." Domínguez-Quintero et al. [79] refer to this as a definition for quality of experience – indeed, it is an appropriate definition for the quality of experience in the context of any commercial service.

Human-computer interaction

It is hard to draw a sharp boundary between the fields of HCI and networking. Classification could be attempted based on the authors and their educational background, terminology used in the paper, or the papers in the list of references. There are also some statistical differences that could be utilized. For instance, acronyms QoS and QoE are more common in networking papers while some authors (e.g., Csikszentmihalyi and Hassenzahl) are more popular in the HCI field. In HCI papers, the perspective usually is the user, whereas in networking papers, even when QoE is an essential concept, the main perspective is often the business of the service provider. However, it is difficult to design a simple rule that provides a reliable classification between these neighboring fields of research. Thus, the classification used here is based on subjective assessment relying on various aspects, including the main research topic and methods used in the paper.

The need for an improved user interface became urgent in the 1980s when personal computers began to become common. The need became even more pressing in the 1990s with the onset of the Internet and web services. Jakob Nielsen's *Usability Engineering* [225], published in 1993, provided an outstanding overview of state of the art in the early 1990s and is still frequently used in usability research. Nielsen did not use the term quality of experience in his book; even experience is relatively rare in Nielsen's texts.

The first notable paper in the HCI field with quality of experience as a key concept was written by Alben in 1996 [9]. Nevertheless, Alben used the concept sparingly: in addition to the title, the term is mentioned only once in the text[17] and a couple of times in illustrations. According to Alben, experience means all aspects related to the use of an interactive product: feelings, understanding, purpose, and context. As to the quality part of the phrase, her main interest is in the work of a design awards jury: successful and engaging experiences are valuable to users and, thus, feasible criteria to be used by the jury.

Regarding the use of the term quality of experience, a paper by Forlizzi & Ford [95] was a key initiator in the HCI field. References in the paper included Alben's paper [9], *Art as Experience* by Dewey [77], and the famous experience economy paper by Pine & Gilmore [237]. One of their contributions was to consider experiences based on narrations and storytelling – although their target audience was interaction designers, a narrative viewpoint could be expedient in other contexts, too.

One year later, Morris & Turner [219] defined quality of experience as a construct that describes how successfully a system can do what is needed functionally. Furthermore, Morris & Turner state that

[17] The only time Alben uses the concept, she puts it into parenthesis "quality of experience." This may indicate some uncertainty as to whether the chosen term is correct or appropriate, but it may well be that Alben uses the parentheses due to the novelty of the term.

their definition resembles Shackel's utility dimension [269] whereas Shackel's likeability and Csikszentmihalyi's flow [64] differ from the quality of experience. A significant merit of [219] is in the broad account of the relationship between utility (or usefulness), usability, quality of experience, and technology acceptance; this is still a critical topic for further studies.

Ramesh Jain [151] argued in 2004 that designers of multimedia systems typically use quality of service as the basis for their design. Jain stressed the need to include users in the process of system design, even though that inclusion requires "soft" subjective measures. Jain took it for granted that readers understood quality of experience without discussing the meaning of the concept and without referring to any previous publication.

McNamara & Kirakowski [203] distinguish three areas of technology usage (experience, usability, and functionality) and argue that these areas should be analyzed separately with their own methodologies. They utilize an extensive set of references to discuss different interpretations of usability and other related concepts. Their insight is that quality of experience was introduced into the HCI field to address those aspects of technology usage that the quality of use or usability perspective largely omits. In this sense, quality of experience is an extension of usability studies towards subjective experiences and emotions.

As to the difference between user experience (UX) – a popular term in HCI – and quality of experience, Wechsung & De Moor [326] provide an illuminating discussion.[18] The appendix in [326] adeptly summarizes the main differences between quality of experience and user experience. For instance, they claim that quality of experience resembles more the concept of customer experience used in business studies than user experience. Moreover, user experience as a field of study is more qualitative and theoretically oriented than those quality of experience studies that are related to communication networks. They argue that QoE and UX originate from QoS and usability and have even to some degree replaced them. It is, however, worth noting that many early HCI papers using the concept of quality of experience (e.g., Alben [9], Forlizzi & Ford [95], McNamara & Kirakowski [203], Morris & Turner [219], and Suri [293]) did not mention quality of service. Thus, it is possible that quality of experience was independently invented in the HCI domain, while the later use of QoE in HCI partly originated from networking research.

One of the key authors in HCI, Marc Hassenzahl, noted that experience becomes *user experience* by focusing on a specific mediator of experiences, namely interactive products [113]. Hassenzahl favored *user experience* over *quality of experience.* Some other authors see the relationship differently. Lallemand et al. [182] mention one possible definition for user experience "The quality of experience a person has when interacting with a specific design." This would mean that quality of experience is the genus[19] for user experience, which is a reasonable choice as the user is one of the many roles of a person. Like Wechsung & de Moor, Lallemand et al. [182] argue based on an extensive expert survey that UX is not considered a marketing concept and has a strong scientific basis.[20]

In HCI, quality of experience is just one of the numerous terms used to discuss and analyze what happens in the user's mind during the interaction with technical devices. User experience and usability likely remain the most prominent terms in the HCI field, whereas quality of experience will typically be used as an auxiliary term to expand the range of user experience towards emotional aspects.

Networking

Quality of Service (QoS) became a popular concept in analysis of communications networks and services in the 1990s. In addition to the standardization of QoS mechanisms for various networking

[18] According to Google Scholar, there are about as many papers with both "user experience" and UX as there are papers with both "quality of experience" and QoE. However, the number of papers using just the phrase "user experience" is about thirteen times higher than the number of papers with the phrase "quality of experience." This report concentrates on the quality of experience papers.

[19] genus = a class of objects divided into several subordinate species.

[20] The difference between UX and QoE studies is also reflected in their respective vocabularies, which are discussed in Appendix 3 of this report.

technologies, hundreds of QoS papers were published every year [174]. In contrast, the earliest networking paper in Set Q3, Hamada et al. [109], was not published until 2000.[21] Hamada et al. used the acronyms QoS and QoE in the context of network management but did not give any source for the acronyms. Besides, due to the limited number of citations, their paper has had only a minor effect on the later quality of experience literature.

As Schatz et al. [262] notice, the first consequential paper promoting quality of experience in the context of networking was written by van Moorsel [312]. Van Moorsel, who then worked at Hewlett-Packard (HP), refers to several reports published by HP Labs and two industrial undertakings, qualityofexperience.org and webqoe.org (HP participated in both organizations). His paper refers to [39],[22] which uses the concept of QoS experience, but not quality of experience. HP's initial concern was the business analysis of web services that could be divided into three parts: QoS, QoE (download completion probability, response time, etc.), and QoBiz (quality of business defined by cost and revenue elements). Van Moorsel had two main messages. First, unlike QoS, QoE has a subjective element. Second, QoE metrics may be influenced by external systems that are not in the direct control of the service provider and, thus, not included in the QoS evaluation.

Some other technical papers using quality of experience [115, 130, 259] were published at the same time (2001-2002) as [312]. Three factors unite these papers. First, they did not specify a source for the concept. Second, the quality of experience was a relatively minor issue. Third, the papers have not had any noticeable effect on the later use of the concept. In any case, they serve as an indication that the concept began to gain popularity in the networking field. As to the early adoption of QoE, the paper by Khirman & Henriksen [171] was more significant because it introduced a method to assess QoE, the cancellation rate of web services as a function of latency and bit rate. The only reference in their paper that contains QoE is a white paper published by webqoe.org.[23]

Siller & Woods [279] provide a long list of references covering essential QoS references and some early QoE papers, like [222], IETF RFCs, and ETSI and ITU-T standards.[24] All the QoE references in [279] are published by private companies: Patricia Seybold Group [10], Empirix [85], InfoLibria [117], Polycom [226], and HP [331]. A paper written by Telenor researchers [120] contains a similar list of references, including [9, 171, 226, 312]. Siller & Woods [279] state that the origin of QoE metrics is what was earlier considered the user perception of QoS. They also propose a definition for QoE: "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."

As to standardization, ITU-T G.1000 [137], published in 2001, did not mention the term quality of experience but defined *QoS perceived by user/customer* as "a statement expressing the level of quality that customers believe they have experienced." The same term (abbreviated as QoSE) and definition appear in ITU-T E.800 [142] published in 2008. The first ITU-T standard defining QoE was ITU-T P.10/G.100 [143] published in 2008. The standard defined quality of experience as "the overall acceptability of an application or service, as perceived subjectively by the end-user." The first IETF document discussing QoE, RFC6390 [58], was published in 2011. The authors of the RFC noted that "QoE may also be considered as QoS delivered, received, and interpreted by a user with the pertinent qualitative factors influencing his/her perception of the service." Thus, QoE could be interpreted as the user's opinion about the quality of the offered service.

[21] A paper by Elwalid et el. [84] published in 1996 mentions quality of experience once but there is no indication that the paper had any effect on the later use of quality of experience.
[22] More precisely, van Moorsel refers to an HP report with the same title and authors as the journal article [39].
[23] Quality of experience appears in the title of a workgroup "The Web Quality of Experience Workgroup" formed by HP, Cisco, and other Internet companies (InfoWorld March 26, 2001, p. 33). I know nothing about the origin of the quality of experience in the title; possible sources include Alben [9] and Pine & Gilmore [237].
[24] ETSI: European Telecommunications Standards Institute, ITU-T: International Telecommunication Union Telecommunication Standardization Sector, IETF: Internet Engineering Task Force, RFC: Request for Comments.

The next commonly cited technical quality of experience papers [173, 330] were published in 2008. As Winkler & Mohandas noted at the beginning of their paper [330], quality of experience had become a common term as regards video and multimedia services. The first year in which the networking area represented the majority of papers in Set Q3C (13 out of 24 papers) was 2008. From 2009 onwards, over 80% of papers using the term have been related to networking technologies and services.[25]

Fiedler et al. [90] is the most cited networking paper in Set Q3 published before 2015.[26] Fiedler et al. offer an extensive discussion about the relationship between QoS and QoE. Their opening statement deserves to be highlighted: "Quality of experience ties together user perception, experience, and expectations to application and network performance, typically expressed by quality of service parameters." This is an apt objective for quality of experience studies, and it can also be interpreted as a definition of quality of experience. The oldest reference in [90] is [41], a paper published in 2000 by HP discussing quality and QoS but without mentioning QoE.

A report with 31 authors from both HCI and networking fields provides a valuable discussion about the meaning of key terms [47]. They define quality of experience as "the degree of delight or annoyance of the user of an application or service." The same definition was later adopted by ITU-T [145] and is now the most common definition of quality of experience. The paper [47] briefly discusses the relationship between user experience and QoE and notes that QoE is also related to the content of the service, not only to the use of the service. Moreover, they observe that while QoS is about system performance, QoE covers a wide range of system features, including users' expectations and socio-economic issues. Regardless of the broad scope and lengthy discussion of QoE, they do not refer to any philosophical, psychological, or business article.

For the past 15 years, the most popular topic for quality of experience research has been video, first under the term IPTV (Internet Protocol television) [118, 170, 176, 294]. These IPTV articles were technical without any discussion about the nature of experiences. The method to assess the quality of IPTV services was usually mean opinion score (MOS). After 2010, the terminology changed to internet video [22] and HTTP (Hypertext Transfer Protocol) video streaming [212]. In recent years, the term has been established in the form of HTTP adaptive streaming (HAS) [101, 155, 229, 266, 340].

In summary, the use of quality of experience as an established concept in the networking field originates from the needs of product vendors. For instance, Hewlett-Packard was actively promoting new web services with this fresh concept. In general, private companies were active in QoE debates from 2001 to 2009, when the meaning of QoE evolved. The number of networking papers in Set Q3C with the first author from a private company was 31% during this period. In contrast, the corresponding share was only 3% from 2015 to 2020. The practical nature of many early technical papers using quality of experience implies that the term was chosen without any awareness of its use in other fields, like philosophy, psychology, and tourism. The main motivation for the use of quality of experience was vendors' need to convince potential customers that they mastered not only technology but also end-user needs and experiences. The term quality of experience was then widely adopted by the networking community due to its positive connotation and the wide area of applicability from technical optimization to business development and marketing.

[25] In Set Q3C, 82% of papers published in 2009 or later are classified as networking papers, while 11% of the papers are related to HCI. Note, however, that the classification between networking and HCI is vague.

[26] There is one networking paper in Set Q3 with more citations, Agiwal et al. [5] with 2255 citations. In [5, table VIII], the authors have selected six references to represent QoE studies. Only two of them [21, 25] belong to Set Q3. Four others include a white paper (Huawei), a public-private partnership project, an article with 8 and another article with 25 citations in Google Scholar (7.4.2022). Although they refer to Alben [9], Chen et al. [56], and Seufert et al. [266], the merits of the paper are more in 5G technology than in human experiences.

Interaction between fields

As the previous sub-sections reveal, quality of experience is a concept that has been used in several scientific fields.[27] Nevertheless, it seems that there has not been much interaction between the fields of philosophy, psychology, tourism, and networking technology. There even is a tendency to form narrower research communities (e.g., voice, video, and data) as observed by Schatz et al. [262]. Multidisciplinary papers may also encounter problems when submitted to specialized journals, which may diminish the number of recognized papers with a wide scope.[28] The formation of research silos may even hinder the progress of scientific studies. Gary Marcus expressed his concern as follows "I'm going to go out on a different limb and say that the biggest obstacle to science in some fields that I could name is dismissive ignorance of literature outside of practitioners' own immediate expertise."[29]

There is hardly anyone who claims that multidisciplinary research is undesirable in the case of quality of experience research. Wechsung & De Moor [326, appendix] even stated that multidisciplinary research is done in the context of quality of experience, increasingly also in practice.[30] They compared the situation with user experience research in which multidisciplinary research has been prevalent from the beginning. Similar ideas have been expressed by Varela et al. [313].

When the main topic of a paper is limited to a technical issue, such as the performance of a video streaming application, it is understandable that the fundamental nature of experiences is left out of discussion. In contrast, when the paper contains conceptual discussion, it is reasonable to make use of a wider range of source material also from other fields of study. In the case of quality of experience, numerous papers have contributed to the conceptual discussion [47, 71, 102, 182, 244, 250, 279, 326, 333, 344]. Some of these papers utilize literature from other fields of study. For instance, Csikszentmihalyi [64] is quoted in [333], Dewey [77] in [326, 344], McCarthy & Wright [201] in [102, 182], and Parasuraman et al. [230] in [244].

Baraković & Skorin-Kapov [28] note that multidisciplinary studies are required for Web QoE analysis. They provide a list of relevant dimensions [28, figure 1]: effectiveness, trust, aesthetics, usability, quality of information, loading time, pleasure, acceptability, satisfaction, and efficiency. Similarly, when addressing and modeling the effects of user expectations, Sackl et al. [258] use the results from socio-psychology and service quality and consumer satisfaction research to assess user expectations in general. Shin [274] emphasizes that the industry should focus more on QoE instead of just monitoring QoS. He proposes the following QoE factors: service, content, hedonicity, coolness, system, and utility. These factors require multiple viewpoints, at least psychology, economics, and technical performance.

Regardless of these positive examples, the interaction between different areas of quality of experience research has remained rather limited. This observation is illustrated in Table 2, which shows the share of papers citing a particular author and the share of different research methods. The eminent philosopher John Dewey has been quoted in philosophical, psychological, and HCI papers, e.g., in [48, 95, 203]. In contrast, Dewey is quoted very rarely in tourism and networking papers. According to Buchanan [48], John Dewey's *Art as experience* [77] was highly influential in developing design practices; HCI is a kind of art. Csikszentmihalyi is a rare example of an author who is regularly cited in a variety of research areas, perhaps because flow is an easy-to-understand and useful term when studying diverse experiences.

[27] The use of one concept in several scientific fields and the importance of conceptual analysis are thoroughly addressed by G. Canguilhem [49].

[28] Note also that the selection process for quality of experience articles in this report was partly based on the number of citations in Google Scholar. Additional effort has been put into finding multidisciplinary quality of experience articles with a limited number of citations.

[29] https://twitter.com/garymarcus, 10.4.2022. The limb Marcus refers to in his tweet is the problem of expensive laboratory equipment that creates a financial barrier for scientific research (Ali Mohebi's tweet, 10.4.2022).

[30] Wechsung & De Moor [326] is the most multidisciplinary paper in Set Q3C when the criterion is the diversity of authors in references. Other multidisciplinary include [111, 173, 182, 295].

The most popular references in the area of tourist and recreational experiences are Baloglu & McCleary [24] and Otto & Ritchie [228] with 5100 and 1613 citations, respectively. Still, no paper in Set Q1 outside tourism and recreation refers to these two papers. Similarly, tourism papers seldom contain any references to HCI or networking papers. Because the fields of HCI and networking partly overlap, the interaction between the fields has been regular, for instance, in the form of the Dagstuhl seminars [72, 91] and the International Conferences on Quality of Multimedia Experience (QoMEX). For instance, out of the 16 QoMEX papers cited in this report, eight are classified in the HCI area and the other eight in the network area. Regardless of this collaboration, one of the most-cited authors in the HCI field, Marc Hassenzahl [113, 114], is cited only in three networking papers in Set Q3C.

Technical papers refer almost exclusively to other technical papers. For example, [6, 25, 124] provide valuable summaries of 5G design requirements and use the acronym QoE systematically but contain purely technical references. To my knowledge, the fascinating paper by Minhas & Fiedler [207] is the only networking paper in Set Q1 mentioning that the term quality of experience has also been used in the fields of philosophy and psychology. The references in [207] include Dewey [78], Csikszentmihalyi [66], and Harman [112]. In Set Q1, the only papers referring to both Csikszentmihalyi and Harman are [173] and [207].

The two papers in Set Q3C [60, 228] that refer to both Csikszentmihalyi and Parasuraman are related to tourism. All services offered to tourists depend entirely on the experiences created by the service provider. Thus, psychology and business analysis are key ingredients in every comprehensive analysis of tourist attractions. Another tourism paper by Ellis et al. [83] offers an extensive analysis of experiences and refers to Dewey's completed experience, Kahneman's fast-thinking (System 1)[31], and Csikszentmihalyi's flow experience.

In the case of QoE, the main authority in technical contexts is ITU-T. Sometimes even standards change their opinions. In ITU-T P.10/G. [143], published in 2008, QoE is defined as "the overall acceptability of an application or service, as perceived subjectively by the end-user" whereas the newer version of the same standard [145], published in 2018, defines QoE as "the degree of delight or annoyance of the user of an application or service." The latter version is dominant in the current technical literature.

[31] As to the dichotomy of System 1 / System 2, see [160].

Table 2: The share of papers referring to certain authors and standards in Set Q3C, - denotes no paper and 0% denotes a case in which 0 < share < 0.5%.

	Philos-ophy	Psychology & education	Tourism & recreation	HCI	Networking	All fields until 2015	All fields 2016-2020
Number of papers	11	73	34	73	480	446	243
Authors mentioned							
J. Dewey	18%	15%	-	5%	-	4%	-
M. Csikszentmihalyi	-	71%	12%	7%	1%	13%	0%
R. M. Ryan	-	40%	-	1%	0%	7%	-
A. Parasuraman	-	-	44%	1%	0%	4%	1%
S. Baloglu	-	-	12%	-	-	1%	0%
J. Otto	-	-	21%	-	-	1%	0%
M. Hassenzahl	-	-	-	11%	1%	2%	-
M. Fiedler	-	-	-	10%	13%	11%	8%
ITU-T	-	1%	-	52%	40%	36%	29%
E.800 [142]	-	-	-	4%	6%	5%	3%
P.10/G.100 [143, 145]	-	-	-	7%	7%	6%	4%

The methods used in different fields are summarized in Table 3. Some of the methods, especially ANOVA (analysis of variance) and RMSE (root-mean-square error), do not indicate much about the nature of the research due to their simplicity. Nevertheless, we can make one curious observation. If we consider only the most recent papers in Set Q1 (438 papers in total), we can divide them into two groups: those belonging to Set Q3C (243 papers) and those not belonging to Set Q3C (195 papers). In the first group, only 8 papers (or 3%) mention ANOVA, whereas in the second group 29 papers (or 15%) mention ANOVA. In contrast, there is not any similar difference in the older papers published in 2015 or earlier. Thus, it seems that the mention of ANOVA in a recent paper reduces its relevance in the context of quality of experience. However, we need to be careful not to overstate this finding, because there are so many keywords and so many possible ways of grouping the papers that it is inevitable to find some statistically significant correlations.

The Experience Sampling Method (ESM) is a common tool in psychological studies [66, 183, 263]. The concept of quality of experience can be defined as one of the factors in the ESM method. Although ESM is rarely used in other fields of research, there are some interesting studies reported in [27, 73, 132].

Table 3: Methods in different fields in Set Q3C. Abbreviations: ESM: Experience Sampling Method, ANOVA: analysis of variance, SEM: Structural Equation Modelling, MOS: mean opinion score, SSIM: Structural Similarity index measure, TAM: technology acceptance model, RMSE: root-mean-square error, JND: just-noticeable difference. Physiological measurements include EEG, ECG, EMG, EDA, NIRS, and fMRI. All percentages are rounded to the nearest whole number, - denotes no paper at all.

	Philosophy	Psychology & education	Tourism & recreation	HCI	Networking	All fields until 2015	All fields 2016-2020
Number of papers	11	73	34	73	480	446	243
Physiological meas.[32]	36%	3%	-	10%	3%	3%	6%
ESM	-	55%	3%	1%	1%	9%	-
ANOVA	-	18%	9%	11%	4%	9%	3%
SEM	-	4%	35%	3%	-	3%	2%
TAM	-	-	-	6%	0%	1%	-
RMSE	-	-	-	11%	6%	5%	7%
SSIM	-	-	-	20%	13%	10%	14%
MOS	-	1%	-	44%	42%	35%	31%
JND	-	-	-	6%	2%	2%	2%
Machine learning							
mentioned	9%	-	-	15%	22%	9%	36%
applied	-	-	-	3%	7%	2%	13%

The most common method used in tourism research is Structural Equation Modeling (SEM) [13]. Tian-Cole et al. [302] was the first paper in Set Q3C in which SEM was used to analyze tourists' behavioral intentions. In their model, quality of experience affects overall satisfaction, which, in turn, affect behavioral intention. Similarly, in the structural model designed by Um et al. [309], perceived quality of service affects satisfaction, which, in turn, affects revisit intention. In their terminology, satisfaction is defined as a quality of experience referring to an emotional state of mind. SEM is also occasionally used in the HCI field. For instance, Shin utilizes SEM to study the nature of experiences in a virtual environment [275] and in the context of Internet of Things [274]. In contrast, no networking paper in Set Q1 has used SEM as the main research method.

The technology acceptance model (TAM) [184] is mainly applied in the HCI field [333]. TAM is used to assess the perceived benefits of the technology compared to the cost of using it. The only networking paper in Set Q1 referring to TAM states that it is used to analyze acceptability before actual usage, whereas QoE research is concerned (according to Schatz et al. [261]) with the acceptability in use. Moreover, one recent paper in Set Q3, Videnovik et al. [315], adopts TAM to analyze the use of augmented reality games in education. Moreover, some technical papers discuss the acceptability without referring to TAM, like [70, 126, 290].

Two other methods, net promoter score (NPS) [126, 193] and just-noticeable difference (JND) [249] expand the methodological variation of QoE studies. Since the main application of NPS is the measurement of customer loyalty, it can be used to evaluate the effect of QoE on customer behavior. The only papers in Set Q1 using NPS, [1, 2], are methodically identical to studies based on the conventional MOS scale. JND is applied in the field of experimental psychology and has a strong scientific foundation. In the QoE context, JND has been utilized as a subjective method like MOS [290, 297, 346] and as a basis

[32] As to physiological methods and machine learning, see Appendix 4.

for constructing objective methods [14, 56, 341]. NPS and JND are examples of methods borrowed from other disciplines, albeit with limited impact on the quality of experience research.

Present state of research

This brief section provides an overview of the current state of quality of experience research. The history of quality of experience and the present state is summarized in Table 4. Based on the papers in Set Q3C, we can argue that the focus of quality of experience research is gradually moved towards specific, typically technical, matters. The most striking change is the decrease in the proportion of articles from other areas than HCI and networking: the proportion of the other fields is only 4% in 2016-2020 compared to 28% before 2016. This change can be explained mainly by the rapid increase of number of technical papers while the number of non-technical quality of experience papers seems to be slowly declining.[33]

As to interaction between fields, of the 446 papers published in 2015 or earlier (in Set Q3C), 18 papers referred to Dewey, and 64 papers referred to Csikszentmihalyi. In the period 2016-2020, no paper in Set Q3C refers to Dewey and only one paper, Sabet et al. [256], refers to Csikszentmihalyi. Nevertheless, some technical papers have a broad scope, but they rarely reach the required citation limit for Set Q3C. For instance, Zhang et al. [345] and Morrisey et al. [220] mention Dewey and refer to McCarthy & Wright [201]. Similarly, Möller et al. [216] use Csikszentmihalyi's concept of flow. None of these three papers [216, 220, 345] reached the citation limit for Set Q3C. Thus, there is no indication that interaction between different fields is becoming more active.

The role of technology is often seen as an efficient way to control everything from the environment to humans. For instance, Barakabitze et al. [26] divide QoE management into three components: monitoring & measurement, optimization & control, and modeling & assessment. In another paper written by Bouraqia et al. [42], providers face the challenges related to delivering, measuring, and controlling QoE. These two papers summarize the technical view on QoE: it is something that service providers shall measure, control, and even optimize on behalf of users and customers. On the contrary, a philosopher or a psychologist may find the idea of controlling the quality of an experience strange or perhaps even dangerous.

The share of authors affiliated with a private company has always been small except in the field of networking, where device vendors and network operators were active at the beginning of the quality of experience research. The share of papers originating from private companies was 31% between 2000 and 2009 whereas a great majority of the current quality of experience research is carried out in universities in all areas.[34]

[33] It is worth noting that the increase in the number of technical papers raises the citation threshold for Set Q3C, which partly explains the decrease in the number of non-technical papers in Set Q3C. However, the number of psychological papers in Set Q3C published between 2005 to 2010 was 26 whereas there is no psychological paper in Set Q3C published in 2015 or later. In contrast, the decline in the number of tourism papers in Set Q3C (from 14 to 5) might be explained by the higher citation threshold due to the avalanche of networking papers.
[34] I have worked at Aalto University for 14 years. However, I was deeply influenced by the pragmatic atmosphere at my former employers, Telecom Finland (1990-1995) and Nokia (1995-2008).

Table 4: The main fields for quality of experience research are illuminated by key publications describing the generic field, a more specific topic, and a recent topic. The numbers in parenthesis: (publication year, the number of citations in Google Scholar)

Generic field	Philosophy	Psychology	Business analysis	Usability	Communication networks
	Dewey, Art as Experience [77] (1934, 20188)	Csikszentmihalyi & LeFevre, Optimal experience in work and leisure [66] (1989, 2840)	Parasuraman et al., A conceptual model of service quality and its implications for future research [230] (1985, 36837)	Nielsen, Usability engineering [225] (1993, 23391)	Xiao & Li, Internet QoS: a big picture [334] (1999, 1456)
Specific topic	**Experience**	**Education**	**Tourism**	**HCI**	**Internet services**
	Harman, The intrinsic quality of experience [112] (1990, 1632)	Schiefele, Interest, learning, and motivation [263] (1991, 2259)	Baloglu, A model of destination image formation [24] (1999, 5100)	Alben, Quality of Experience: Defining the criteria for effective inter-action design [9] (1996, 603)	van Moorsel, Metrics for the internet age: Quality of experience and quality of business [312] (2001, 267)
Recent topic	**Consciousness**	**Positive technology**	**Recreation**	**User experience**	**Adaptive streaming**
	Tononi & Koch, Consciousness: here, there and everywhere? [304] (2015, 562)	Riva et al., Positive technology: using interactive technologies to promote positive functioning [252] (2012, 332)	Jin et al. (2015). The effect of experience quality on perceived value, satisfaction, image and behavioral intention of water park patrons [156] (2015, 337)	Lallemand et al., User experience: A concept without consensus? [182] (2015, 251)	Seufert et al., A survey on quality of experience of HTTP adaptive streaming [266] (2015, 820)

2 Framework and definitions

As the previous sections have shown, a rich literature on various aspects of quality of experience is available for all interested readers and researchers. Furthermore, there are numerous methods to measure and analyze experiences in different contexts. Nevertheless, much fewer papers systematically consider experiences in a way that covers several fields of study. This section aims to fill this gap by introducing a framework that is applicable independent of the specific characteristics of any research field.

Experience

The word experience has an ancient origin as thoroughly reviewed by Martin Jay in his inspirational book *Songs of Experience* [154],[35] but here we proceed straight to the current use of experience. Oxford English Dictionary[36] gives the following definition: experience is "The fact of being consciously the subject of a state or condition, or of being consciously affected by an event. Also an instance of this; a state or condition viewed subjectively; an event by which one is affected." A conceivable interpretation for *one* in the last sentence is the conscious mind. As John Dewey has stated, the mind is a verb [77, p. 274]. [37] Thus, the role of the mind is more active than just being affected by external stimuli.

What are the elements of the activities occurring in the mind? Ernest Hilgard [122] argues that traditionally the mind has been thought of as a trilogy of cognition, affection, and conation.[38] A similar structure can be found in Baloglu [23] in which attitude is conceptualized as having three parts: cognition (perceptions/beliefs), affect (feelings), and conation (behavioral intentions).[39] Similarly, Raake & Egger state that experiences have emotional, sensory, conceptual, and actional features [244]. Because conscious refers to a mind and being affected can be called a process, experience could be defined as the process in which the mind interacts with the environment by perceiving, feeling, thinking, and acting.[40]

Are the four elements (perceiving, feeling, thinking, and acting) sufficient to cover all important aspects of experiences? Moreover, are all four elements truly necessary? Finally, would it be possible to choose other similar but more feasible terms? In my opinion, feeling and thinking are absolutely necessary for describing experiences.[41] Moreover, in the context of users experiencing communication services, perception is an integral element that cannot be omitted. Thus, acting is the most debatable of the four aspects.

One may claim that instead of acting, an experience can only contain behavioral intention, which is a regularly used variable in structural equation models, e.g., [53, 303].[42] In this report, the approach is based on the conviction that acting is a part of the same process that comprises perceiving, feeling, and

[35] I recommend Jay's *Songs of Experience* [154] for everyone interested in the etymology and history of the concept of experience.
[36] https://www.oed.com/
[37] There is a long history of philosophical discussion about whether the mind is passive or active. See, for instance, the section *Locke and the experience of the senses* in [154, p. 44-56].
[38] Conation: an inclination (such as an instinct, a drive, a wish, or a craving) to act purposefully, https://www.merriam-webster.com/dictionary/conation.
[39] According to Bakker [20, Table 5], already Plato mentioned the triad of feeling, thinking, and acting. The same list of verbs is proposed by Thompson [300, p. 16].
[40] See also Walls et al. [318] who present a valuable list of definitions for experience in the context of customer experience.
[41] Emotion could be used instead of feeling. Similarly, thinking could be replaced by cognition and perceiving by perception.
[42] The logic in [303] is the following: programs, amenities, entertainment ⇒ overall experience ⇒ overall satisfaction ⇒ re-visit intention. In this structure, satisfaction is a separate construct from experience.

thinking. It is, therefore, not enough to consider only the planning of an action or the intention to act, but the action itself. This is obviously the case during flow experiences.[43]

Moreover, some authors define quality of experience based solely on thinking or acting. For instance, Daniel Dennett has explained consciousness based on his multiple drafts model where the mind primarily is a cognitive information processing unit [75]. In contrast, B. F. Skinner has said that "to experience the world is to test it" [281, p. 80]. Nonetheless, these two approaches are not utilized in this report due to their relatively low popularity in the current scientific research.

Finally, it is also possible to define some additional elements of experience. Possible additions include *imagination* and *memory* [154, p. 246][44], *intuition* [305], and *motivation* [326].[45] The most relevant of the possible additions is, in my opinion, Berndt Schmitt's [264] concept of *relate* referring to "social-identity experiences that result from relating to a reference group or culture." The other four experiential modules in [264] are essentially the same that I am proposing here (sense, feel, think, and act). However, I leave the topic of social aspect of experiences for further studies.

To facilitate the further examination of quality of experience, an appropriate framework would be beneficial. We have as a basis the four parts of the experience, perceiving, feeling, thinking, and acting. In addition, there are external factors affecting experiences and consequences from the actions made during the experiences. Consequently, we can draw a simple framework describing the main aspects of experiences as illustrated in Figure 1.

Figure 1 also shows six options to define the quality of experience. In the first option, quality of experience is measured and defined by the accuracy of perception. In the second option, quality of experience refers to satisfaction. In the third option, experience refers to all feelings during an experience. In the fourth option, experience covers perceiving and feeling, while the more active parts of experience (thinking and acting) are omitted. In the fifth option, experience contains (in addition to the four basic elements) the influencing factors generated by needs, incentives, and the environment and the main immediate effects of the experience. In the sixth option, experience extends to the whole area depicted in Figure 1.

In a summary, we can define six ranges for studying and defining quality of experience:

R1: perceiving,

R2: satisfaction,

R3: feelings,

R4: subjective experience consisting of perceptions and feelings,

R5: momentary experience including inputs and outputs, and

R6: all aspects of life.

[43] For me, the strongest source of flow is dancing. During the most gratifying dancing experiences, perceiving, feeling, and acting are entirely intermingled while occasional moments of conscious thinking have only minor effects on the overall experience.

[44] The modes of experience that include imagination and memory were originally proposed by R. D. Laing in *The Politics of Experience and The Bird of Paradise* (the term used by Laing is "modality of experience").

[45] Wechsung & de Moor [326] refer to Marc Hassenzahl when mentioning motivation.

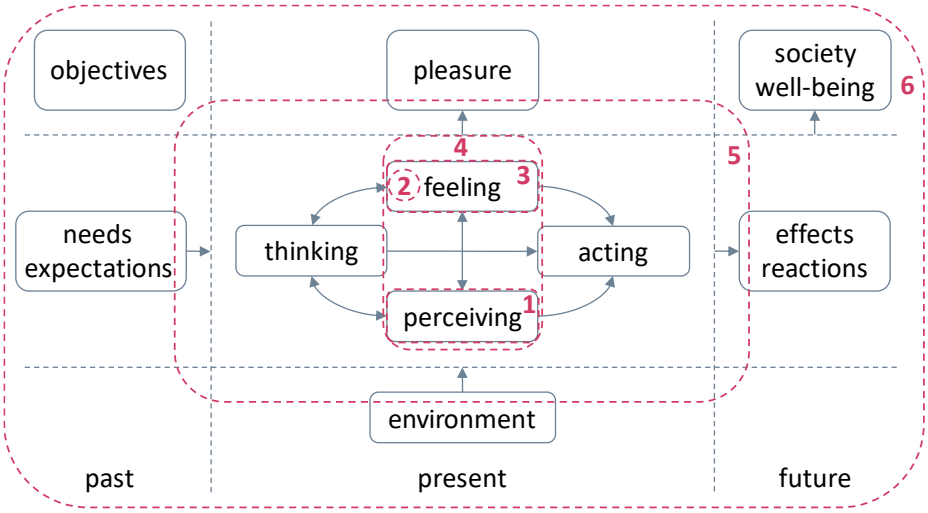

Figure 1: The past, present, and future of an experience with six ranges (marked by dotted circles): R1: perceiving, R2: satisfaction, R3: feelings, R4: subjective experience consisting of perceptions and feelings, R5: momentary experience including inputs and outputs, and R6: all aspects of life.

Quality

Reeves & Bednar [248] note in their outstanding article that quality has been defined as value, conformance to specifications, conformance to requirements, fitness for use, loss avoidance, and meeting and/or exceeding customers' expectations. In this list of options, quality is a property of commercial products, which is an important aspect in many quality of experience studies, but not the only possible viewpoint. Thus, let us continue with a dictionary definition. Oxford English Dictionary defines quality as follows: "Originally: the nature, kind, or character (of something). Later: the standard or nature of something as measured against other things of a similar kind; the degree of excellence possessed by a thing." Regardless of the long history of the concept, it is still unclear, at least in the case of quality of experience, whether quality should mean

E1: the excellence of a thing (usually high)

E2: the degree of excellence of a thing (low or high), or

E3: any characteristics of a thing.

Moreover, the use of an abstract abbreviation, like QoE, can blur or bias the meaning of the concept. For instance, the following expressions are relatively common in networking literature: "applications have specific QoE needs [12]" and "to ensure high QoE for each application [25]." Nevertheless, hardly anyone explicitly claims that applications have needs or experiences.[46] In these and numerous other papers, quality seems to refer to the properties of the service instead of the properties of experiences. Another typical claim in networking papers is that QoE measures the service quality as perceived by the user [30, 56, 158, 266]. According to this definition the object of optimization is the service quality, not the experience quality. In contrast, many researchers in the tourism sector explicitly distinguish service quality from experience quality: service quality refers to the quality of service controlled by the supplier, while experience quality refers to the psychological effects of services [11]. Clear separation between these two concepts would be beneficial also in the case of network services.

[46] At least if need is defined as "a physiological or psychological requirement for the well-being of an organism" https://www.merriam-webster.com/dictionary/need

Another possible term that can be used in the same was as quality of experience is *qualia*.[47] It is mainly used in the philosophical literature, see, e.g. [304].[48] Qualia refer to the purely subjective nature of experience related to perceiving and feeling, particularly to those aspects "which no amount of purely physical information can convey" [147]. Qualia could, therefore, be used when only the subjective aspects of experience are embraced. A somewhat similar term is *quality of experiencing* proposed by Raake & Egger [244] but that concept has not gained a foothold in quality of experience research.[49]

In summary, it is possible to distinguish the following four depths of quality of experience analysis:

D1: the quality of a service or product,

D2: the specific experiences created when a specific service or product is used,

D3: the overall nature of experience during a specific event, and

D4: the fundamental nature of conscious experience.

However, these two dimensions, excellence (E) and depth of assessment (D) are not independent. In practice, E1 is typically used with D1 (for marketing purposes), E2 is used regularly with D1 and D2 and sometimes also with D3, whereas E3 is mostly used with D3 and D4. Thus, we primarily use ranges (R1-R6) and depths (D1-D4) in the following discussion. However, it is important also to keep in mind the excellence dimension when we later discuss the metrics used for quality of experience.

Table 5 provides an overview of the definitions for quality of experience given by different authors in different fields of study.

[47] Formally, "quale" is the singular form while "qualia" is the plural form, https://plato.stanford.edu/entries/qualia/

[48] Raake & Egger [244] is the only technical (networking or HCI) paper in Set Q3C using the concept qualia.

[49] Quality of experiencing is sometimes used in educational research. For instance, in [277] "concrete experience" and "abstract conceptualization" are the two ends of quality-of-experiencing continuum. The main objective of the scale is to describe different learning attitudes, but the scale can be applicable also in other contexts. Raake & Egger [244] is the only article in Set Q3C using the concept quality of experiencing.

Table 5: Quality of experience definitions classified on two dimensions.[50]

Field	Author [reference]	Range (R1-R6)	Depth (D1-D4)
Philosophy	Tononi & Koch [304]	R4	D4
	Gallagher & Zahavi [98]	R4	D4
Psychology	Csikszentmihalyi & LeFevre [66]	R4	D3
	Riva et al. [252]	R6	D3
Tourism	Otto & Ritchie [228]	R3	D2
	Tian-Cole et al. [302]	R4	D2
HCI	Alben [9]	R5	D3
	Wu et al. [333]	R5	D2
	Wechsung & De Moor [326]		
	user experience	R4	D2
	quality of experience	R1	D1
	Shin [276]	R5	D2
Networking	van Moorsel [312]	R2	D1
	Siller & Woods [279]	R1	D1
	Winkler & Mohandas [330]	R1	D1
	Brunnström et al. [47] and ITU-T [145]	R2	D2
	Chen et al. [56]	R1	D1

Chen et al. [56] offer an extreme viewpoint in which the quality explicitly refers to the service: "Quality of experience (QoE) is the perceptual quality of service (QoS) from the users' perspective." As a kind of opposite extreme, quality may refer to the fundamental nature of all experiences – or to "the peculiar diaphanous quality of experience" as dubbed by G. E. Moore.[51] In a way, Moore's intriguing remark may give an explanation why it is so common to consider the service as the object of quality assessment even though the term is the quality of *experience*. Whenever we attempt to consider an experience per se, we easily move on to look at the objects of the experience instead of the experience itself.

Definitions and methods

Now we are ready to compile a set of definitions for quality of experience. The aim is to cover all fields of study and all ranges (R1-R6) defined above. The situation is illustrated in Figure 2.

[50] The classifications are based on my interpretation of the ideas presented in each paper; the authors' intentions and opinions may differ from the presented classification.
[51] See also the item Gallagher & Zahavi in Appendix 1 of this report (non-fiction books).

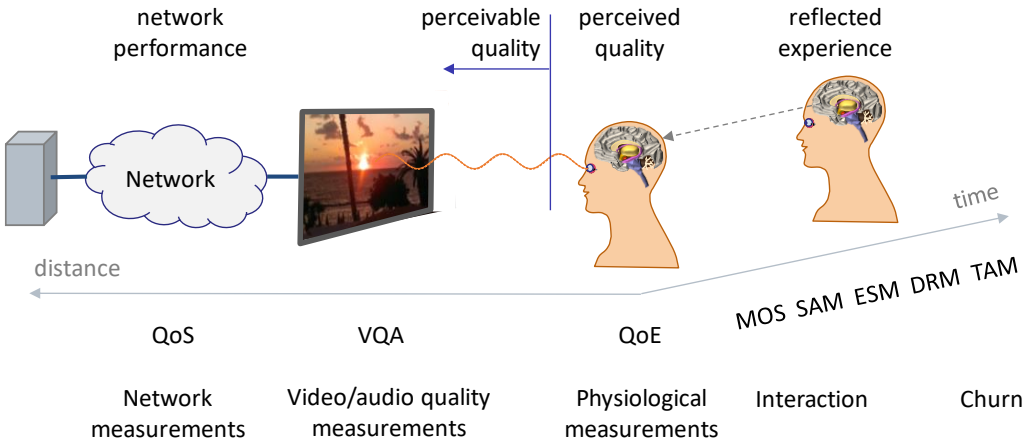

Figure 2: The process of creating, experiencing, and assessing the quality of a video session. VQA: video quality assessment, SAM: Self-Assessment Manikin, ESM: Experience Sampling Method, DRM: Day Reconstruction Method, TAM: technology acceptance model. Churn is a rate of change that occurs in a business over a period of time as existing customers are lost and new customers are added.[52]

QoE as the perceived quality of service

When a service is directly related to sensory perception, such as hearing and seeing, it is reasonable to think that QoE is equal to the quality of the sensory perception. A video stream can emulate the original stream more or less accurately (Figure 2 illustrates the situation). This accuracy can be objectively measured by different methods, for instance, by structural similarity index (SSIM) or peak signal-to-noise ratio (PSNR) (the most commonly used methods are summarized in Appendix 4). The result of objective measurement is sometimes called QoE when the outcome of the method has a sufficiently strong correlation with the results of subjective evaluations in similar situations. Then one may assert, as the authors of [336] do, that PSNR × SSIM can be used as an evaluation metric for QoE. Similarly in [349], QoE is measured on an SSIM scale from 0 to 1. Bentaleb et al. [37] use SSIM and a normalized QoE scale from 1 to 5 but without mentioning the mean opinion score. In these kinds of approaches, the following definition is feasible:

> **Definition 1.** The accuracy of an audio, video, or another sensory flow perceived by the user when compared to the original sensory flow.

Nevertheless, perception is a limited perspective on the complex reality of experiences. The term experience suggests that humans are somehow involved in the analysis even if the measurement has been carried out by a technical system without any direct human involvement. A more suitable term than QoE would be *perceivable quality* where quality clearly refers to the quality of the service or product under consideration. Another possibility, proposed by Akhtar & Falk [7], is to use the term Quality of Perception (QoP). However, QoP is problematic in the sense that we seldom are interested in the quality of the perception process but either in the quality of received content or in the quality of the ensuing experience.[53]

[52] https://www.merriam-webster.com/dictionary/churn
[53] QoP studies may, for instance, evaluate how well people with different backgrounds and capabilities understand spoken English or discern traffic signs.

QoE as a measure of satisfaction

The use of QoE is often limited in the evaluation of user or customer satisfaction. In this case, it is reasonable to use the ITU-T definition for QoE [145]:

Definition 2. The degree of delight or annoyance of the user of an application or service.

The method associated with definition 2 is almost always mean opinion score (MOS) [136]. MOS is used in one-third of the papers in Set Q3C, mostly in technical papers, whereas only two non-technical papers in Set Q3C mention MOS. In the case of MOS research, satisfaction is usually asked immediately after a session. If the object of assessment is a short video or audio clip, it is reasonable to assume that the feeling of satisfaction/dissatisfaction is still present and does not require arduous conscious reflection.

QoE as a measure of feelings

When the diversity of devices and applications to be evaluated evolves and expands, it becomes important to analyze effects other than pure satisfaction with the current products. For example, immediate satisfaction is an insufficient method for assessing the true value of a health-related service. A service event may create hope or anxiety, love or frustration, or a spectrum of emotions that cannot be condensed on a simple scale describing the level of satisfaction. Similarly, all autonomous devices, like an autonomous car, create numerous emotions that affect the actions to be taken, and satisfaction is only a small part of the big picture. Thus, the following definition could be applicable when an event creates a variety of feelings:

Definition 3. The spectrum of emotions that a product or service creates in the customer.

The variety of emotions is so large that it is hardly ever reasonable to encompass all of them.[54] Thus, any practical research must apply a simplified model. In psychology, the two main dimensions to assess emotions are 1) affect or hedonic valence (happy, cheerful, sociable, satisfied, etc.) and 2) potency or arousal (active, alert, strong, excited, etc.) [66, 195]. However, these dimensions are not necessarily orthogonal. For instance, when ESM is used to measure emotions, there often is a significant correlation between valence and arousal (0.60-0.74 in [66]). A three-dimensional system of feelings typically consists of pleasure, arousal, and dominance [20].

Self-Assessment Manikin (SAM) [44, 245] is a promising method to efficiently assess several emotional dimensions in parallel. Gupta et al. [107] use a two-dimensional (valence-arousal) SAM-scale and EEG (Electroencephalography) to evaluate the perception of speech quality. Their results indicate that MOS and Valence/SAM produce similar results. Beyer et al. [38] apply a three-dimensional SAM and EEG to measure the effects of video quality in cloud gaming. Raheel et al. [245] utilize a similar setting (three-dimensional SAM and EEG) to analyze the experiences related to tactile enhanced multimedia. Although SAM has been used in user experience research [326], Raheel et al. [245] is the only paper in Set Q3C applying SAM.

Subjective quality of experience

Descartes said, "*Cogito, ergo sum.*"[55] This expression attaches great importance to thinking compared to other aspects of experience. The prevalent tradition in neuroscience considers the brain primarily as a thinking machine. For instance, Matthew Cobb acknowledges that emotion has been barely touched in his book about the brain [59, p. 383]; feeling and experience do not even appear in the index of the book. This contrasts with the analysis of experience in technical fields that rarely mentions thinking as an integral part of the experience. To alleviate this conundrum, it is possible to state that quality of experience refers to subjective events in the mind, not only to feelings and emotions. However, if we

[54] See, e.g., Wikipedia's article *Emotion classification*, https://en.wikipedia.org/wiki/Emotion_classification.
[55] I think, therefore I am.

limit the scope to purely subjective aspects, we may need to exclude thinking and acting, because they are more directly observable (through speaking, writing, and acting) than feeling and perception. So we end up with the next definition:

> **Definition 4.** The intrinsic nature of the purely subjective processes taking place in the brain.

However, subjective processes related to an experience are hard to measure and observe, even by the person who has or undergoes the experience. We are mostly unaware of the convoluted processes that lead to diverse perceptions and feelings.[56] In that sense, the conscious mind does not have control over satisfaction.[57] Thinking can affect satisfaction by drawing attention to certain aspects of event. From a measuring viewpoint, that is an advantage, as satisfaction can considered mostly (though not fully) independent of conscious thoughts. Any deeper analysis of a purely subjective process also requires philosophical reflection.

Quality of experience as an interactive process

In all previous options, the experience was presumed to be an internal process taking place in the brain. In this fifth option, however, the range of experience is expanded to the immediate antecedents and the direct effects of an experience. Laghari et al. [181] define QoE "as a blueprint of all human subjective and objective quality needs and experiences arising from the interaction of a person with technology and with business entities in a particular context." Wu et al. [333] define QoE as "a multi-dimensional construct of perceptions and behaviors of a user, which represents his/her emotional, cognitive, and behavioral responses, both subjective and objective while using a system." Similarly, the following definition emphases the role of interaction:

> **Definition 5.** The nature of the process where a person interacts with the environment by perceiving, feeling, thinking, and acting

This definition indicates that quality of experience is partly determined by the environment and is always a result of complex interactions to the extent that interaction can be considered an integral part of the experience. This definition is approaching the concept of user experience (see [326]).

Although most HCI and networking papers assume that QoE can be measured on the MOS scale in a controlled laboratory setting, many papers acknowledge that in reality, many psychological and contextual factors affect the experience of a user. For example, Song et al. [284] state that in addition to measuring the quality of the delivered service, an appropriate analysis shall take into account the user's needs and desires. This crucial issue has been discussed in general in [163, 254, 342] and in a technical context in [71, 257, 258, 351]. A multitude of papers using the concept of influencing factors is available, see [47, 129, 188, 214, 251, 266, 346, 348]. Baraković Husić et al. [29] provide an outstanding summary of influence factor studies and perceptual dimensions.

It should be noted that the introduction of influencing factors does not necessarily imply an extension to the range of quality of experience itself. On the contrary, the factors can remain external to the core of quality of experience, which generally remains the level of satisfaction with regard to a service.

As an example of a study embracing social interactions, Botella et al. [40] give an overview of positive technologies. They utilize a set of three scales to assess human functioning: hedonic, eudaimonic, and social scales. The hedonic scale refers to feelings and immediate value, the eudaimonic scale to self-realization and resilience, and the social scale to the ability to interact with other people. The challenge is to integrate the certainly important social aspect into the predominantly individual analysis of quality of experience.

[56] See, for instance, Barrett [31], Cobb [59], and Prinz [243].
[57] We are not necessarily satisfied with our immediate reactions. For example, see https://www.gocomics.com/pearlsbeforeswine/2016/05/02.

In another example, Tsiropoulou et al. [306] present a game-theoretical model in which museum visitors maximize their QoE. Shah-Mansouri & Wong [270] provide another example of game-theoretical analysis. Both papers presume that customers are selfish and attempt to maximize their own (short-term) utility. The fundamental assumptions about human behavior are critical in the future when people are even more intermingled through new technologies; altruistic behavioral patterns can create positive qualities of experience.

Quality of experience as a measure of success or well-being

Finally, quality of experience can refer to an approach where human experience plays a key role in analyzing the effects of all kinds of events, products, and services on the success of a large organization or the well-being of society.[58] Then, the main perspective is that of an external stakeholder instead of the experiencing person. In this case, we can devise the following definition:

> **Definition 6.** The degree to which a conscious mind contributes to the well-being or success of a stakeholder.

The stakeholder can be either a business actor (a service provider or a product vendor) or a social system, like family, community, society, or even humankind. For example, Riva et al. [252] propose "to use technology to manipulate the quality of experience, to increase wellness, and generating strengths and resilience in individuals, organizations, and society." There are only a few sociological papers in Set Q1 and even fewer relevant for this report. The most significant example is Csikszentmihalyi's insightful paper titled *Leisure and socialization* [63].

Similar to user experience studies, current QoE research focuses on individual experiences while the social aspect is often overlooked; Georgopoulos et al. [103] provides a rare example in which the authors briefly discuss the harmful effects of selfish maximization of QoE. In contrast, many happiness studies indicate that the well-being of a person depends more on social relationships than anything else. For instance, Bruno Frey [96, p. 151-152] lists and gives weights for ten ways to improve personal happiness. Three of the ways are social: provide help to others (weight = 1.5), make friends and value them (2.5), and get married (3). In contrast, earning money has a weight of 0.5 while there is no indication that buying any new product or service would create noticeable happiness.[59] There is a need to expand the analysis, if we are seriously bothered with human well-being, particularly in the case of social media applications.

Any thorough business and societal analysis requires a multidisciplinary approach that includes engineering, HCI, psychology, sociology, economics, and even philosophy. Hammer et al. [111] provide valuable discussion about the challenges when the analysis embraces eudaimonic aspects like meaningfulness, self-actualization, autonomy, competence, and relatedness. The scope of [111] is the most extensive among all the quality of experience papers in Set Q1. However, they prefer the term quality-of-user-experience instead of quality of experience. In general, it might be better to use some other term than quality of experience to model the effects of experiences on the total well-being of society.

[58] A possible approach is to follow Derek Parfit's ideas, see [231]. Parfit notes that self-interest theory is neutral with respect to time but partial with respect to person whereas present-aim theory is partial to both persons and time. Most QoE analyses fall into the latter category. A truly moral standpoint should consider not only the current well-being of other people but also the well-being of future generations. Certainly, this is a strict requirement for quality of experience research, but in view of the looming climate crisis it is unacceptable to only consider the immediate enjoyment of a video and to ignore the consumption of non-renewable resources.
[59] In contrast, using money to buy time can promote happiness [328].

Naming the definitions

As the previous chapters reveal, there is a wide variety in the definitions related to quality of experience. However, it is questionable to apply the same label, quality of experience, to all six definitions. Thus, the approach here is to use different labels for different definitions.

My proposal is that the form quality of experience (without capitalization) should be reserved for a broad sense that covers also other aspects of experiences than pure satisfaction with a service or product. The fifth definition is the best for this purpose. In contrast, the second definition standardized by ITU-T shall use the abbreviation QoE and the capitalized form Quality of Experience. The most limited, first definition should not be called quality of experience, because the object of quality analysis is not an experience but a product – thus, a more suitable label is perceivable quality (of a product or service).

If the business analysis is wider than the satisfaction analysis, the term experience quality is applicable, see, for instance, the analysis of festival attributes by Tian-Cole et al. [303]. Then if the research topic is the fundamental nature of subjective experience (fourth definition), qualia is a promising candidate. Finally, in the case where the objective is to analyze the effects of individual experiences on the profitability of a business actor or the well-being of society, my recommendation is to use the term utility, or more specifically terms experienced utility and decision utility. Table 6 provides a summary of the different definitions and terms.

Table 6: Summary of proposed definitions and terms (the first column refers to the numbering in Figure 1).

	Definition	Scope	Proposed term	Methods
1.	The accuracy of an audio, video, or another sensory flow perceived by the user when compared to the original sensory flow.	optimization of multimedia apps	perceivable quality	SSIM, PSNR, etc.
2.	The degree of delight or annoyance of the user of an application or service.	technical & business optimization	**QoE**	MOS
3.	The spectrum of emotions that a product or service creates in the customer.	far-reaching business development	experience quality	SEM
4.	The intrinsic nature of the purely subjective process taking place in the brain.	physiology / philosophy	qualia	philosophy
5.	The nature of the process where a person interacts with the environment by perceiving, feeling, thinking, and acting.	psychology	**quality of experience**	ESM, DRM
6.	The degree to which a conscious mind contributes to the well-being or success of a stakeholder, e.g., the society.	well-being and business analysis	utility	surveys

3 Mean opinion score

History

ITU-T published a recommendation for the speech transmission quality in 1984 [135]. The specific question given in the questionnaire was: "Which of these four words comes closest to describing the quality of the *connection* during conversation?" Possible answers were excellent, good, fair, and poor. The question was likely designed by a network engineer instead of a usability designer or a business developer because the object of the question was not related to the experience during a phone call or the overall quality of telephone service but the performance of a technical system.

In [135, annex A], the choices were numbered from 1 (excellent) to 4 (poor). However, it seems that the numbers were used to design the questionnaire rather than to illustrate the differences between the choices. In the treatment of the results section, it is noticed that the choices should be accorded scores from 4 (excellent) to 1 (poor) to calculate the mean opinion score (MOS).

The MOS-scale was used already in the 1960s. According to Duncanson [80] "Excellent, Good, Fair, etc." was perhaps the most popular set of terms to assess telephone transmission quality. Duncanson also presents other sets of terms and even gives a numerical value for each term on a scale from 0 to 1 as follows: Excellent 1.00, Good 0.81, Fair 0.51, Poor 0.32, Bad 0.24, and Unacceptable 0.11.[60]

When the variety of services began to expand from voice to video and multimedia in the 1990s, there was a long tradition to measure the technical quality of communications services using a scale with four or five stages and to use the mean opinion score as a basis of further analysis. For example, [97] published in 1997, applies MOS to assess video quality and offers "a systematic method to estimate the required bandwidth to guarantee user's preference on video quality."[61] This tradition of using MOS led to three important consequences. First, the viewpoint was technical as users were asked to assess the connection or transmission quality, not the quality of their experiences. Second, the mean opinion score inevitably required a numerical scale, while most of the papers did not include any discussion about the suitability of the scale for statistical analysis. Third, the main research goal was to control network resources – and sometimes also users – instead of analyzing experiences.

This was the situation during the first years of the millennium when the term quality of experience became more common in the contexts of HCI and network services. It is worth noting that MOS became common in quality of experience literature only somewhat later; the oldest papers in Set Q3C using MOS [170, 188] were published in 2006. Subsequently, MOS was quickly adopted as the main method to assess user opinions. In the period from 2007 to 2011, 62% of networking papers and 50% of HCI papers in Set Q3C mentioned MOS. Several papers provide valuable summaries of subjective and objective methods using MOS as a metric for measuring and modeling QoE, see [45, 126, 180, 282, 290, 307].

Problems with MOS

Regardless of its popularity, MOS has several weaknesses related to the naming of the labels, the range of the scale, and the intrinsic nature of the scale. According to Pinson & Wolf [238] and Sullivan et al. [291], MOS values from different subjective experiments are not directly comparable, even when they are carried out in similar laboratory settings. Watson & Sasse [324] address the labeling problem of opinion scales; it is almost impossible to translate a label to another language without changing the properties of the scale. Another problem with the typical MOS scale is that the established use of a scale from 1 to 5 does not cover the whole range of possible outcomes. ITU-T [146, figure 2] recommends

[60] The numerical values are based on Thurstone's Law of Categorical Judgment [301]. As far as I can judge, this type of numerical scale should not be interpreted as a utility-scale. Instead, the primary purpose of the terms (excellent, bad, good, etc.) is to facilitate communication with other people, and that purpose determines the cognitive distance between the terms. See also Appendix 5.

[61] Nowadays, the end of the sentence would probably be written: "to guarantee quality of experience."

using, in the context of game experience, an extended version of the MOS scale with extremely bad (below bad) and ideal (above excellent), but without elaborating the properties of the scale. Duncanson [80] also proposed (in 1968) an extended scale with *unacceptable* below *bad*. Furthermore, another scale presented in [80] defines the lowest quality level as *conversation impossible*. Something similar could be useful also in the case of novel communication services.

The conversion of terms (excellent, good, etc.) to a numerical scale is an even more critical problem. Most of the terms describing service quality can be usually put in preference order: good is preferable to fair, fair is preferable to poor, etc. Moreover, it is easy to assume that the difference between two consecutive terms is constant. For instance, most of the QoE models reviewed by Barman & Martini [30, equations 1-8] seem to handle MOS as a scale with a constant distance between consecutive levels.[62]

The nature of the MOS scale is mentioned in the ITU-T standard [136], in which the annex E includes a short notice that MOS should be assumed as an ordinal scale in data analysis. Similarly, Mitra et al. [208] state that arithmetic operations cannot be applied to subjective ratings such as MOS. To solve this dilemma, Seufert [267] proposes to use the full QoE distribution instead of simple (average) MOS. The full distribution can then be analyzed by multinominal distributions. Seufert also emphasizes that MOS is a scale with categorical values and that ratios between categorical values are not meaningful [267]. Xu et al. [337] referred to the problems with MOS when discussing utility-based models and argued that averaging shall be applied on the utility-scale instead of the conventional MOS scale. Hoßfeld et al. [128] recognized the same problem but with a differing solution in which the analysis should consider the standard deviation of the opinion score in addition to the mean opinion score. Other quality of experience papers discussing this critical issue include [54, 73, 152, 291].

To clarify the nature of the challenges with MOS, we can utilize the theory of scales proposed by S. S. Stevens in 1946 [288]. The theory contains four types of scale:

- A **nominal** scale of measurement is used to assign objects into discrete categories. The nominal scale allows the determination of equality (=, ≠) but no other mathematical operation.
- An **ordinal** scale of measurement is used to assign values to objects based on their relative ranking with respect to one another. It allows the determination of greater or less (<, >) but not more complex mathematical operations.
- An **interval** scale of measurement is used to assign values to objects in such a way that numerically equal distances on the scale represent the equal distances between the properties of the measured objects. It allows the determination of equality of intervals and, thus, addition and subtraction (+, -).
- A **ratio** scale is an interval scale having a fixed and meaningful zero value. It allows the determination of equality of ratios and, thus, multiplication and division (x, /).

Now the question is, how should we classify the ordinary MOS scale consisting of five levels? There are three different aspects to be considered. First, there is a *cognitive distance* between the terms, e.g., between excellent and good.[63] Everyone with a satisfactory knowledge of English likely agrees that excellent is better than good, but exactly how much, is a debatable issue. Moreover, one may argue that the cognitive distance between excellent and good is larger than the cognitive distance between poor and bad, but there is hardly any consensus on this issue. Thus, the MOS scale is ordinal, but whether it is an interval scale is arguable.[64]

Second, the effect of perceived service quality on the subsequent *behavior* of the user is a separate issue from the cognitive distance. In other words, we cannot assume that a constant cognitive distance means a constant difference in any objectively measurable behavior. For instance, if a customer changes

[62] See: xkcd, Assigning numbers, https://xkcd.com/2610/.
[63] Jones & McManus [157] and Watson & Sasse [324] used the term "distance." Chen et al. [54] introduced the term "cognitive distance" while referring to [324], which, in turn, refers to [157].
[64] The nature of the cognitive scale is further considered in Appendix 5.

her opinion about service quality from excellent to good, the change can cause a small increase in the probability of changing the service provider. On the contrary, when the perceived quality is changed from fair to poor, the addition to the same probability likely is much larger.

Third, a service event, whether it is a phone call, a video streaming session, or a network game, affects the *well-being* of the user. It seems natural to assume that an excellent service would have a positive effect on the user's well-being whereas a bad service experience would have a negative effect on the user's well-being. If the service quality is good enough for all practical needs, the effect likely is minor while if the service quality is unsatisfactory there could be a noticeable negative effect on the user's well-being. It should be, however, remembered that the real source of well-being seldom is the technical quality of service, but the content of the event, e.g., rewarding conversation with a friend, excitement when playing a game, or all various feelings created by a movie. In reality, we try to avoid considering technical issues during an event unless the discomfort due to poor quality exceeds a threshold.

The main point underlined here is that when MOS values are used for different purposes, it is not enough to convert the average values from one scale to another scale, but the conversion must include the whole distribution of MOS values. For instance, one may aim to optimize the total experience among users sharing certain network resources by optimizing the use of resources between different video streams. Is it enough to maximize the average MOS over all users? The answer definitely is no for well-being and behavioral scales. Thus, the average MOS gives a valid answer to the question:[65] "What is the average opinion about service quality?" On the contrary, this average MOS does not give valid answers to questions like: "What is the effect of an intervention that reallocates the network resources on the total well-being of all users?" or "What is the effect of service quality on the customer decisions to change the service provider?" The three scales are illustrated in Figure 3.

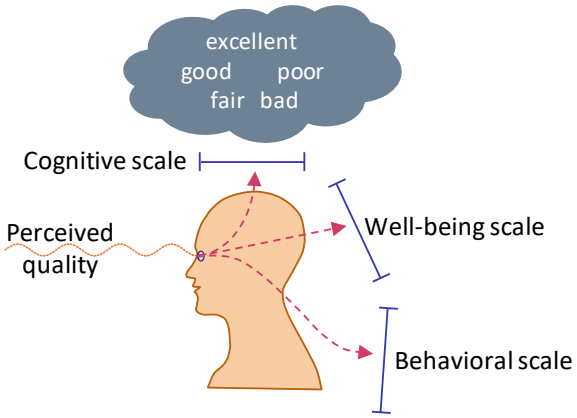

Figure 3: Three scales to assess the quality of service: cognitive scale, well-being scale, and behavioral scale.

Thus, it is obvious that these three numerical scales (cognitive distance, well-being, and behavioral) are essentially different.[66] In the cognitive distance scale, the difference between excellent and good is significant while the difference between poor and bad may be somewhat smaller; the reverse seems to be true with well-being and behavioral scales. Consequently, MOS cannot be used as an interval scale *without determining the intended use of the scale*.

[65] Average MOS refers to opinions given by different users with *different levels* of service quality, not mean opinions of different users experiencing identical service events.

[66] I was not aware of this difference before writing this section in April 2022. The term "cognitive distance" played a key role in the process of improved understanding; it is hard to be aware of something without a proper name. To cite Georg Simmel [280, p. 179]: "Our intellect can grasp reality only as a modification of pure concepts, which, no matter how much they diverge from reality, are legitimized by the service they render in the interpretation of reality."

Further studies

So far, there is no established model to convert MOS values to other scales to allow the use of statistical methods. In general, any optimization of well-being or customer behavior should be conducted on a properly defined scale instead of the MOS scale. An appropriate utility-scale might even be interpreted as a ratio scale with a well-defined zero level. A possible approach on a behavioral scale would be (see Figure 4 and Table 7):

1) Zero means that there is no noticeable effect on the behavior of the user or customer compared to the expected situation.

This scale is typically negative in the sense that it is used to measure the experienced annoyance compared to reasonable expectations and its effects on ensuing decisions. The effects may be measured on a ratio scale where "the number of disturbances during video sessions" affects "the amount of annoyance" which, in turn, affects "the likelihood of leaving the service." From the service provider's viewpoint, it is reasonable to add up the likelihoods of losing a customer. The same scale can also be used when great experiences encourage a customer to praise the product to friends, which may, in turn, increase the number of paying customers.

On the well-being scale we may make the following assumptions (see Figure 4 and Table 7):

2) Zero refers to a situation that is a total waste of time (but no more harmful than that). For example, the quality of a voice call can be so bad that conversation is impossible. In that case, a forced continuation of the call would mean a total waste of time.

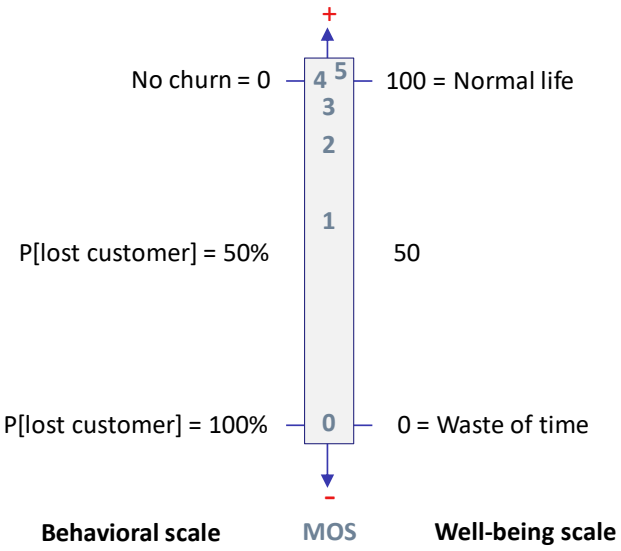

Figure 4: A tentative relationship between behavioral scale (left), mean opinion score (middle), and well-being scale. The locations of MOS values are for illustration purposes. [67]

[67] Figure 4 presents only one MOS scale in the middle of the figure, but that is not a statement that the conversions from MOS to behavioral scale and from MOS to well-being scale were identical.

Table 7: Scales used for different purposes.

Name of the scale	Metrics	Object of measurement	Zero point	Unit	Type of scale
Cognitive scale	Extended MOS (0 – 5)	Service quality	Useless service	1 step	Ordinal, interval (?)
Well-being scale	Experienced utility	Value of time or life	Totally wasted time	Value of normal life	Interval, ratio (?)
Behavioral scale	Decision utility (negative)	Probability of leaving service	No effect	Probability of losing the customer	Ratio
	Decision utility (positive)	New customers through word of mouth	No effect	New customer	Ratio

In the case of the behavioral scale, for the service provider, two options are presumed to be equally (un)attractive: 1) one customer is surely lost while another one surely remains a customer, and 2) each of the two customers is lost with a probability of 50%. Since a service provider's revenue (or profit) is an interval scale, it makes sense to convert each customer's satisfaction onto the revenue (or profit) scale, and then do further calculations on the monetary scale. The middle column contains numbers from 0 to 5. The locations of the numbers are my guesses about the location of MOS values relative to the behavioral scale (thus, the locations of the MOS numbers are for illustrating purposes).

The nature of the well-being scale is less clear. We may, however, assume that there is a point (50) where a customer is equally likely to choose between two events: 1) a two-minute event with well-being of 50, and 2) an event with one minute of normal life (100) and another minute with a total waste of time (0). This property of the well-being scale is reasonable assuming that a person's overall well-being can be estimated simply by integrating momentary well-being values over a longer period. However, this hypothesis is by no means obvious due to various psychological phenomena.

Although I (with my colleagues) have constructed well-being models based on the value of time with a zero level [175, 240], I am still unsure whether the value of time scale could be used as a ratio scale. To claim that the value of time can be improved, say, 5 percent through specific improvements of service quality is debatable.

Many articles analyzing the performance of adaptive streaming begin with measurable quality parameters (start-up delay, average bit rate, video quality, etc.) with the goal of calculating a parameter called QoE. One example of such an approach is formula 5 in [340]:

$$QoE_1^K = \sum_{k=1}^{K} q(R_k) - \lambda \sum_{k=1}^{K-1} |q(R_{k+1}) - q(R_k)| - \mu \sum_{k=1}^{K} \left(\frac{d_k(R_k)}{C_k} - B_k \right)_+ - \mu_s T_s$$

where K is the number of components, C is the bit rate, R is chunk size, B is the playout buffer when the chuck download started, and T is the startup delay. Although it can be assumed that maximizing this technical QoE parameter improves user satisfaction, it is questionable to call the parameter quality of experience. Nonetheless, the authors in [340] state "It also achieves significant improvement (60+% median QoE) compared to the industry reference player dash.js."[68] Likewise, authors in [196] state that a method provides "QoE improvements ranging from 12%-25%." Despite these statements, the relationship between the QoE improvements and subjective experiences remains unclear, because it is

[68] "It" at the beginning of the sentence refers to MPC (Model Predictive Control) approach introduced and evaluated in [340].

difficult to imagine how an experience can be improved 12% or 60%.[69] This problem has also been addressed by Michael Seufert in a recent article [268].[70]

Another issue related to the MOS scale is whether the quality to be assessed should be compared with a standard or whether the assessment can be a stand-alone process that does not require any benchmark. It seems natural to assume that satisfaction depends on expectations that can vary between people and change over time. Thus, it seems reasonable to explicitly set a benchmark that is used to assess the quality of experience with a given service. ITU-R has, for instance, defined the double-stimulus continuous quality-scale (DSCQS) method, in which the assessor is asked to view a pair of pictures, one under examination and one directly from the source [134]. DSCQS method has been used in many QoE papers [246, 271, 319, 320].

4 Recommendations

The content of this report can be condensed into the following seven recommendations.

1. **QoE, Quality of Experience, and quality of experience**

   When quality of experience is used in the narrow sense of user or customer satisfaction, and particularly when the MOS scale is used as a metric, it is recommendable to use the acronym QoE and the form Quality of Experience (instead of quality of experience).

2. **Perceivable quality**

   The result of an objective measurement of perceptual accuracy should not be called QoE. A more appropriate term for the result of a perceptual analysis is perceivable quality (or perceivable quality of service).

3. **The nature of the MOS**

   An ordinary MOS scale must not be used as a ratio scale. However, MOS can be used to measure the cognitive distance between different terms, like excellent and good or bad and poor. If the labels are selected carefully and then numbered, e.g., from 0 to 5, the scale could be close to an interval scale.[71]

4. **Conversion between scales**

   It is important to perform any technical or business optimization using a carefully designed scale. The scale must be reasonable in the sense that a rational person attempts to maximize the expected utility when making decisions (decision utility) or when assessing her overall well-being (experienced utility).

5. **Social aspects of experiences**

   A fundamental assumption in most QoE studies is that people are selfish and, thus, seek to maximize their own momentary well-being or utility. However, well-being is a highly social phenomenon. As the results shown in Appendix 6 indicate, the quality of experiences depends on the social context, whether a person is alone or with other people. Thus, I

[69] Similarly, one may devise a formula: happiness(i,j) = c_1·salary(i,j) – c_2·taxes(i,j) – c_3·compulsory expenses(i,j), where i = person, j = month. Certainly, happiness may depend on salaries, etc., but to call the result happiness and then to state that happiness can be increased by 10% by an intervention would be dubious. The strange percent values for QoE are repeated in a survey article [42] without any criticism.

[70] Authors from technical and economic fields seem to feel a considerable pressure to present numerical results.

[71] Actually, "mean opinion score" means exactly this, an average of cognitive evaluations. Averaging is an acceptable operation when the object of assessment is the same while only the evaluators vary.

strongly recommend including the social aspect in quality of experience studies, even when the purpose is to optimize a technical system.[72]

6. Business analysis

When the research objective is to develop a service business, the model and analysis must involve different stakeholders with conflicting goals. Quality of experience is always part of the human side of the model and should be clearly separated from monetary calculations.

7. Well-being analysis

If the research objective is to improve the well-being of a social system, the analysis must include, in addition to immediate gratification, various aspects of human life, most notably the meaning created by the experience, the practical utility of the event for the person involved, and the external effects on the well-being of other people. I would also recommend using a more diverse range of emotions than pure satisfaction, especially in the case of new technologies like embedded and virtual reality and sophisticated healthcare applications.

I would like to highlight one more thing that is easy to ignore but crucial to our future. The objectification of human experience may be necessary to evaluate and optimize the use of technology, but at the same time, the process supports the ever-deepening role of technology and commercial services in our daily lives. Technologies and services provide a constant stream of pleasurable experiences while alienating us from other, more human aspects of life. As Steve Talbott [296, p. 14] has expressed: "Technology is our hope if we can accept it as our enemy, but as our friend it will destroy us." I hope that everyone, including engineers and economists, will seriously consider the contrasting roles of modern technologies.[73]

By the same token, Ralph Stacey [287, p. 76] warns us against the risks of modeling human behavior: "As soon as one draws a conceptual boundary around particular human interactions and regards them as a system with a life of its own, one objectifies that human interaction." and "Building macro models of human interactions, therefore, inevitably loses the quality of human freedom." Furthermore, George Simmel has offered a long treatise on the intriguing subject of objectifying in his remarkable book [280, p. 136]: "Money objectifies the external activities of the subject which are represented in general by economic transactions, and money has therefore developed as its content the most objective practices, the most logical, purely mathematical norms, the absolute freedom from everything personal."[74] My main point is to remind all researchers that users and customers should not be viewed as passive consumers but complex agents with their dreams and aims as well as aversions and anxieties.

[72] Battarbee, Forlizzi, and Koskinen emphasized the importance of other people in the creation of experiences and introduced the term *co-experience* twenty years ago [32, 33, 94]. However, only a few QoE studies have utilized the concept of co-experience. Shin [275] is perhaps the best example, albeit without quoting the co-experience authors mentioned above. Nevertheless, the social aspect is a key element in the analysis of the tourist experience, see, e.g., Huang & Hsu [131].

[73] I have asked the students of a bachelor-level course (Principles of Information Technology) whether they have attempted to limit using their smartphones or playing computer games. The figures are: 53%, 65%, and 61% of the students have attempted to limit their smartphone usage, in 2020, 2021, and 2022, respectively. Similarly, 28%, 31%, and 38% of students have attempted to limit playing computer games in 2020, 2021, and 2022, respectively. The estimated time spent with a smartphone was around 3 hours and the estimated time playing computer games was about 1 hour and 20 minutes.

[74] As a consequence of this concern, I dislike the use of a monetary scale to model the value of human life – regardless of the fact that I have regularly used it in the value of time models.

5 Conclusions

This report consists of three main parts: a historical account of the use of quality of experience in different fields of study, a framework to systematize the definition(s) of quality of experience, and a discussion on the nature of mean opinion score (MOS) scale. In addition, seven appendixes deal with various special topics related to the quality of experience.

Quality of experience has a long history. For nearly a century, from John Dewey's philosophical thoughts in 1887 to Stephen Crites' religious reflections in 1971, the concept was used occasionally. Psychologist Mihaly Csikszentmihalyi was the first to develop a systematic method of assessing the quality of experience in the 1980s. In the 1990s, the concept was adopted in the tourism sector. The concept took off at the turn of the millennium when the fields of HCI and network services adopted it. Currently, a large majority of quality of experience articles are related to communications networks and services.

The framework introduced in the second part of the report considers quality of experience from different perspectives that lead to different definitions of the key concepts. In philosophy and psychology, quality of experience has a broad meaning that covers several aspects of experience, including perceiving, feeling, thinking, and acting. In contrast, Quality of Experience (QoE) in the context of communication networks has a narrower meaning, the degree of delight or annoyance of the user of an application or service. Based on the discussion, six different terms are proposed to cover the full range of applications and objectives in different fields of study.

In the case of communications services, MOS is the prevalent way of measuring quality of experience. The main strengths of MOS are the standardized measurement procedure, the established scale, and the availability of a large number of research results. The main problems of MOS are related to the use of the scale in statistical analysis. Because MOS is not a ratio scale, the individual MOS values must be converted to a suitable scale to allow mathematical analysis. The required conversion depends on the objective of the study, for instance, optimizing the utility of video services, developing the operator's business, or analyzing the overall well-being of a group of people.

For further studies, it would be beneficial to define a generic scale to assess the value of human experiences. A commonly accepted scale could provide considerable synergistic benefits for all parties in different areas of study. A crucial goal is to incorporate meaning, self-actualization, and collaboration with other people in the quality of experience analysis because, in the long run, those are more important goals than immediate gratification. Likewise, the quality of this report is not determined by the immediate feelings it evokes, but by how it can broaden readers' thinking in terms of services, qualities, and experiences.

References

1. Aazam, M., Harras, K. A., & Zeadally, S. (2019). Fog computing for 5G tactile industrial Internet of Things: QoE-aware resource allocation model. *IEEE Transactions on Industrial Informatics*, 15(5), 3085-3092.
2. Aazam, M., St-Hilaire, M., Lung, C. H., & Lambadaris, I. (2016). MeFoRE: QoE based resource estimation at Fog to enhance QoS in IoT. In: *23rd International Conference on Telecommunications* (ICT), 1-5.
3. AccepTV (2015), Video Quality Monitor, https://www.acceptv.com/en/products_vqm.php (date of access: 18.4.2022, the oldest reference in Google Scholar: 2015).
4. Aggarwal, V., Halepovic, E., Pang, J., Venkataraman, S., & Yan, H. (2014). Prometheus: Toward quality-of-experience estimation for mobile apps from passive network measurements. In: *Proceedings of the 15th Workshop on Mobile Computing Systems and Applications*, 1-6.
5. Agiwal, M., Roy, A., & Saxena, N. (2016). Next generation 5G wireless networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 18(3), 1617-1655.
6. Agyapong, P. K., Iwamura, M., Staehle, D., Kiess, W., & Benjebbour, A. (2014). Design considerations for a 5G network architecture. *IEEE Communications Magazine*, 52(11), 65-75.
7. Akhtar, Z. & Falk, T. H. (2017). Audio-visual multimedia quality assessment: A comprehensive survey. *IEEE Access*, 5, 21090-21117.
8. Akhtar, Z., Siddique, K., Rattani, A., Lutfi, S. L., & Falk, T. H. (2019). Why is multimedia quality of experience assessment a challenging problem? I*EEE Access*, 7, 117897-117915.
9. Alben, L. (1996). Quality of experience: defining the criteria for effective interaction design. *interactions*, 3(3), 11-15.
10. Aldrich, S. E., Marks, R. T., Frey, M. M., Goulde, M. A., Lewis, J. M., & Seybold, P. B. (2000). What kind of the total customer experience does your e-business deliver? Patricia Seybold Group.
11. Altunel, M. C. & Erkurt, B. (2015). Cultural tourism in Istanbul: The mediation effect of tourist experience and satisfaction on the relationship between involvement and recommendation intention. *Journal of Destination Marketing & Management*, 4(4), 213-221.
12. Amjad, M., Rehmani, M. H., & Mao, S. (2018). Wireless multimedia cognitive radio networks: A comprehensive survey. *IEEE Communications Surveys & Tutorials*, 20(2), 1056-1103.
13. Anderson, J. C. & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3), 411-423.
14. Antonakoglou, K., Xu, X., Steinbach, E., Mahmoodi, T., & Dohler, M. (2018). Toward haptic communications over the 5G tactile Internet. *IEEE Communications Surveys & Tutorials*, 20(4), 3034-3059.
15. Anwar, M. S., Wang, J., Khan, W., Ullah, A., Ahmad, S., & Fei, Z. (2020). Subjective QoE of 360-degree virtual reality videos and machine learning predictions. *IEEE Access*, 8, 148084-148099.
16. Arimond, G. (1986). Beliefs, Attitudes, and Intended Behavior of Minnesota Deer Bowhunters. Doctoral dissertation, The University of Minnesota, USA.
17. Arndt, S., Brunnström, K., Cheng, E., Engelke, U., Möller, S., & Antons, J. N. (2016). Review on using physiology in quality of experience. *Electronic Imaging*, 2016(16), 1-9.
18. Arthur, W. B. (2009). *The Nature of Technology: What it is and how it evolves*. Simon and Schuster.
19. Asakawa, K. (2004). Flow experience and autotelic personality in Japanese college students: How do they experience challenges in daily life? *Journal of Happiness Studies*, 5(2), 123-154.
20. Bakker, I., Van der Voordt, T., Vink, P., & De Boon, J. (2014). Pleasure, arousal, dominance: Mehrabian and Russell revisited. *Current Psychology*, 33(3), 405-421.
21. Balachandran, A., Sekar, V., Akella, A., Seshan, S., Stoica, I., & Zhang, H. (2012). A quest for an internet video quality-of-experience metric. In: *Proceedings of the 11th ACM workshop on hot topics in networks*, 97-102.

22. Balachandran, A., Sekar, V., Akella, A., Seshan, S., Stoica, I., & Zhang, H. (2013). Developing a predictive model of quality of experience for internet video. *ACM SIGCOMM Computer Communication Review*, 43(4), 339-350.

23. Baloglu, S. (1998). An empirical investigation of attitude theory for tourist destinations: A comparison of visitors and nonvisitors. *Journal of Hospitality & Tourism Research*, 22(3), 211-224.

24. Baloglu, S. & McCleary, K. W. (1999). A model of destination image formation. *Annals of Tourism Research*, 26(4), 868-897.

25. Bangerter, B., Talwar, S., Arefi, R., & Stewart, K. (2014). Networks and devices for the 5G era. *IEEE Communications Magazine*, 52(2), 90-96.

26. Barakabitze, A. A., Barman, N., Ahmad, A., Zadtootaghaj, S., Sun, L., Martini, M. G., & Atzori, L. (2019). QoE management of multimedia streaming services in future networks: a tutorial and survey. *IEEE Communications Surveys & Tutorials*, 22(1), 526-565.

27. Baraković, S. & Skorin-Kapov, L. (2013). Survey and challenges of QoE management issues in wireless networks. *Journal of Computer Networks and Communications*, 1-28.

28. Baraković, S. & Skorin-Kapov, L. (2017). Survey of research on Quality of Experience modelling for web browsing. *Quality and User Experience*, 2(1), 1-31.

29. Baraković Husić, J., Baraković, S., Cero, E., Slamnik, N., Oćuz, M., Dedović, A., & Zupčić, O. (2020). Quality of experience for unified communications: A survey. *International Journal of Network Management*, 30(3), e2083.

30. Barman, N. & Martini, M. G. (2019). QoE modeling for HTTP adaptive video streaming–a survey and open challenges. *IEEE Access*, 7, 30831-30859.

31. Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Pan Macmillan.

32. Battarbee, K. (2003). Defining co-experience. In: *Proceedings of the 2003 International Conference on Designing Pleasurable Products and Interfaces*, 109-113.

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

34. Baucells, M. & Sarin, R. (2012). *Engineering happiness – a new approach for building a joyful life*. University of California Press.

35. Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of general psychology*, 5(4), 323-370.

36. Beerends, J. G. & Stemerdink, J. A. (1994). A perceptual speech-quality measure based on a psychoacoustic sound representation. *Journal of the Audio Engineering Society*, 42(3), 115-123.

37. Bentaleb, A., Begen, A. C., Zimmermann, R., & Harous, S. (2017). SDNHAS: An SDN-enabled architecture to optimize QoE in HTTP adaptive streaming. *IEEE Transactions on Multimedia*, 19(10), 2136-2151.

38. Beyer, J., Varbelow, R., Antons, J. N., & Möller, S. (2015). Using electroencephalography and subjective self-assessment to measure the influence of quality variations in cloud gaming. In: *Seventh International Workshop on Quality of Multimedia Experience* (QoMEX), 1-6.

39. Bhatti, N., Bouch, A., & Kuchinsky, A. (2000). Integrating user-perceived quality into web server design. *Computer Networks*, 33(1-6), 1-16.

40. Botella, C., Riva, G., Gaggioli, A., Wiederhold, B. K., Alcaniz, M., & Banos, R. M. (2012). The present and future of positive technologies. *Cyberpsychology, Behavior, and Social Networking*, 15(2), 78-84.

41. Bouch, A., Kuchinsky, A., & Bhatti, N. (2000). Quality is in the eye of the beholder: Meeting users' requirements for internet quality of service. In: *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, 297-304.

42. Bouraqia, K., Sabir, E., Sadik, M., & Ladid, L. (2020). Quality of experience for streaming services: measurements, challenges and insights. *IEEE Access*, 8, 13341-13361.

43. Boz, E., Finley, B., Oulasvirta, A., Kilkki, K., & Manner, J. (2019). Mobile QoE prediction in the field. *Pervasive and Mobile Computing*, 59, 101039.

44. Bradley, M. M. & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49-59.

45. Brooks, P. & Hestnes, B. (2010). User measures of quality of experience: Why being objective and quantitative is important. *IEEE Network*, 24(2), 8-13.

46. Brown, P. J. (1989). Quality in recreation experience. In: *Outdoor recreation benchmark 1988: Proceedings of the National Outdoor Recreation Forum*, 412-421.

47. Brunnström, K., Beker, S. A., De Moor, K., Dooms, A., Egger, S., Garcia, M. N., ... & Zgank, A. (2013). Qualinet white paper on definitions of quality of experience. In: *Proceedings of the 5th Qualinet Meeting*, 1-24.

48. Buchanan, R. (2015). Worlds in the making: design, management, and the reform of organizational culture. *She Ji: The Journal of Design, Economics, and Innovation*, 1(1), 5-21.

49. Canguilhem, G. (2001). The living and its milieu (translation by J. Savage). *Grey Room*, (3), 7-31.

50. Casas, P., D'Alconzo, A., Wamser, F., Seufert, M., Gardlo, B., Schwind, A., ... & Schatz, R. (2017). Predicting QoE in cellular networks using machine learning and in-smartphone measurements. In: *Ninth International Conference on Quality of Multimedia Experience* (QoMEX),1-6.

51. Chandler, D. M. & Hemami, S. S. (2007). VSNR: A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Transactions on Image Processing*, 16(9), 2284-2298.

52. Chang, T. Y. & Horng, S. C. (2010). Conceptualizing and measuring experience quality: the customer's perspective. *The Service Industries Journal*, 30(14), 2401-2419.

53. Chen, C. F. & Chen, F. S. (2010). Experience quality, perceived value, satisfaction and behavioral intentions for heritage tourists. *Tourism Management*, 31(1), 29-35.

54. Chen, K. T., Wu, C. C., Chang, Y. C., & Lei, C. L. (2009). A crowdsourceable QoE evaluation framework for multimedia content. In: *Proceedings of the 17th ACM international conference on Multimedia*, 491-500.

55. Chen, M., Ma, Y., Li, Y., Wu, D., Zhang, Y., & Youn, C. H. (2017). Wearable 2.0: Enabling human-cloud integration in next generation healthcare systems. *IEEE Communications Magazine*, 55(1), 54-61.

56. Chen, Y., Wu, K., & Zhang, Q. (2014). From QoS to QoE: A tutorial on video quality assessment. *IEEE Communications Surveys & Tutorials*, 17(2), 1126-1165.

57. Chikkerur, S., Sundaram, V., Reisslein, M., & Karam, L. J. (2011). Objective video quality assessment methods: A classification, review, and performance comparison. *IEEE Transactions on Broadcasting*, 57(2), 165-182.

58. Clark, A. & Claise, B. (2011). Guidelines for Considering New Performance Metric Development. IETF, RFC 6390.

59. Cobb, M. (2020). *The idea of the brain: The past and future of neuroscience*. Hachette UK.

60. Costa, G., Glinia, E., Goudas, M., & Antoniou, P. (2004). Recreational services in resort hotels: Customer satisfaction aspects. *Journal of Sport & Tourism*, 9(2), 117-126.

61. Crites, S. (1971). The narrative quality of experience. *Journal of the American academy of religion*, 39(3), 291-311.

62. Crompton, J. L. & Love, L. L. (1995). The predictive validity of alternative approaches to evaluating quality of a festival. *Journal of Travel Research*, 34(1), 11-24.

63. Csikszentmihalyi, M. (1981). Leisure and socialization. *Social forces*, 332-340.

64. Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*, New York: Harper Perennial.

65. Csikszentmihalyi, M. (2000). The costs and benefits of consuming. *Journal of Consumer Research*, 27(2), 267-272.

66. Csikszentmihalyi, M. & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of personality and social psychology*, 56(5), 815-822.

67. Csikszentmihalyi, M. & Schiefele, U. (1992). Arts education, human development, and the quality of experience. In: Reimer, B. & Smith, B.A. (Eds.), *Arts in education. Ninety-first yearbook of the National Society for the Study of Education*, 169-191.

68. Da Hora, N. D., Asrese, A. S., Christophides, V., Teixeira, R., & Rossi, D. (2018). Narrowing the gap between QoS metrics and Web QoE using Above-the-fold metrics. In: *International Conference on Passive and Active Network Measurement*, 31-43.

69. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319-340.

70. De Koning, T. C. M., Veldhoven, P., Knoche, H., & Kooij, R. E. (2007). Of MOS and men: bridging the gap between objective and subjective quality measurements in mobile TV. In: *Multimedia on Mobile Devices 2007*.

71. De Marez, L. & De Moor, K. (2007). The challenge of user-and QoE-centric research and product development in today's ICT environment. *Observatorio (OBS*)*, 1(3), 1-22.

72. De Moor, K., Fiedler, M., Reichl, P., & Varela, M. (2015). Quality of Experience: From Assessment to Application. Report from Dagstuhl Seminar 15022.

73. De Moor, K., Ketyko, I., Joseph, W., Deryckere, T., De Marez, L., Martens, L., & Verleye, G. (2010). Proposed framework for evaluating quality of experience in a mobile, testbed-oriented living lab setting. *Mobile Networks and Applications*, 15(3), 378-391.

74. De Sousa, R. (1990). *The Rationality of Emotion*. MIT Press.

75. Dennett, D. C. & Kinsbourne, M. (1992). Time and the observer: The where and when of consciousness in the brain. *Behavioral and Brain Sciences*, 15(2), 183-201.

76. Dewey, J. (1887). Knowledge as Idealization. *Mind*, 12(8), 382-396.

77. Dewey, J. (1934). *Art as Experience*, New York: Perigee Books.

78. Dewey, J. (1938). *Experience and Education*. Collier Macmillan Publishers.

79. Domínguez-Quintero, A. M., González-Rodríguez, M. R., & Paddison, B. (2020). The mediating role of experience quality on authenticity and satisfaction in the context of cultural-heritage tourism. *Current Issues in Tourism*, 23(2), 248-260.

80. Duncanson, J. P. (1968). Measurement of user opinion of telephone transmission quality. In: *Proceedings of the 4th International Symposium on Human Factors in Telephony*, VDE Verlag GmbH, Berlin, 47-62.

81. Egan, D., Brennan, S., Barrett, J., Qiao, Y., Timmerer, C., & Murray, N. (2016). An evaluation of Heart Rate and ElectroDermal Activity as an objective QoE evaluation method for immersive virtual reality environments. In: *Eighth International Conference on Quality of Multimedia Experience* (QoMEX), 1-6.

82. Egger, S., Hoßfeld, T., Schatz, R., & Fiedler, M. (2012). Waiting times in quality of experience for web based services. In: *Fourth International Workshop on Quality of Multimedia Experience*, 86-96.

83. Ellis, G. D., Freeman, P. A., Jamal, T., & Jiang, J. (2019). A theory of structured experience. *Annals of Leisure Research*, 22(1), 97-118.

84. Elwalid, A. I., Freundlich, G. G., Gerhardt, P. M., Hagirahim, H., Ramakrishnan, K. G., & Tse, D. (1997). An overview of the multimedia communications exchange (mmcx) and its performance characterization. *Bell Labs Technical Journal*, 2(2), 15-35.

85. Empirix (2001). Assuring QoE on Next Generation Networks. Whitepaper, Empirix.

86. Engelke, U., Darcy, D. P., Mulliken, G. H., Bosse, S., Martini, M. G., Arndt, S., ... & Brunnström, K. (2016). Psychophysiology-based QoE assessment: A survey. *IEEE Journal of Selected Topics in Signal Processing*, 11(1), 6-21.

87. Falk, T. H., Pomerantz, Y., Laghari, K., Möller, S., & Chau, T. (2012). Preliminary findings on image preference characterization based on neurophysiological signal analysis: Towards objective QoE modeling. In: *Fourth International Workshop on Quality of Multimedia Experience*, 146-147.

88. Farrel, A. (2009). IANA Considerations for Three Letter Acronyms. *IETF*, RFC 5513.

89. Fen, S. N. (1948). A transactional conception of experience as art. *The Journal of Philosophy*, 45(26), 712-718.

90. Fiedler, M., Hoßfeld, T., & Tran-Gia, P. (2010). A generic quantitative relationship between quality of experience and quality of service. *IEEE Network*, 24(2), 36-41.

91. Fiedler, M., Kilkki, K., & Reichl, P. (2009). From quality of service to quality of experience, Executive Summary. Dagstuhl Seminar 09192.

92. Finley, B. & Kilkki, K. (2014). Exploring empirical rank-frequency distributions longitudinally through a simple stochastic process. *PloS One*, 9(4), e94920.

93. Fisher, R. A. (1921). Studies in crop variation. I. An examination of the yield of dressed grain from Broadbalk. *The Journal of Agricultural Science*, 11(2), 107-135.

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

95. Forlizzi, J. & Ford, S. (2000). The building blocks of experience: an early framework for interaction designers. In: *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*, 419-423.

96. Frey, B. S. (2010). *Happiness: A revolution in economics*. MIT Press.

97. Fukuda, K., Wakamiya, N., Murata, M., & Miyahara, H. (1997). QoS mapping between user's preference and bandwidth control for video transport. In: *Building QoS into Distributed Systems*, 291-302.

98. Gallagher, S. & Zahavi, D. (2012). *The Phenomenological Mind (2nd ed.)*. Routledge.

99. Gallarza, M. G. & Saura, I. G. (2006). Value dimensions, perceived value, satisfaction and loyalty: an investigation of university students' travel behaviour. Tourism management, 27(3), 437-452.

100. García, B., Gallego, M., Gortázar, F., & Bertolino, A. (2019). Understanding and estimating quality of experience in WebRTC applications. *Computing*, 101(11), 1585-1607.

101. Garcia, M. N., De Simone, F., Tavakoli, S., Staelens, N., Egger, S., Brunnström, K., & Raake, A. (2014). Quality of experience and HTTP adaptive streaming: A review of subjective studies. In: *Sixth International Workshop on Quality of Multimedia Experience* (QoMEX), 141-146.

102. Geerts, D., De Moor, K., Ketyko, I., Jacobs, A., Van den Bergh, J., Joseph, W., ... & De Marez, L. (2010). Linking an integrated framework with appropriate methods for measuring QoE. In: *Second International Workshop on Quality of Multimedia Experience* (QoMEX), 158-163.

103. Georgopoulos, P., Elkhatib, Y., Broadbent, M., Mu, M., & Race, N. (2013). Towards network-wide QoE fairness using OpenFlow-assisted adaptive video streaming. In: *Proceedings of the 2013 ACM SIGCOMM workshop on Future human-centric multimedia networking*, 15-20.

104. Glynn, I. (2003). *An Anatomy of Thought: The Origin and Machinery of the Mind*. Oxford University Press, USA.

105. Gosseries, O., Di, H., Laureys, S., & Boly, M. (2014). Measuring consciousness in severely damaged brains. *Annual Review of Neuroscience*, 37, 457-478.

106. Grassini, S. & Laumann, K. (2020). Questionnaire measures and physiological correlates of presence: A systematic review. *Frontiers in Psychology*, 11, 349, 1-21.

107. Gupta, R., Laghari, K., Banville, H., & Falk, T. H. (2016). Using affective brain-computer interfaces to characterize human influential factors for speech quality-of-experience perception modelling. *Human-centric Computing and Information Sciences*, 6(1), 1-19.

108. Haas, L. F. (2003). Hans Berger (1873–1941), Richard Caton (1842–1926), and Electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry*, 74(1), 9.

109. Hamada, T., Czezowski, P., & Chujo, T. (2000). Policy-based management for enterprise and carrier IP networking. *Fujitsu Sci Tech Journal* 36(2), 128-39.

110. Hameed, S., Badii, A., & Cullen, A. J. (2008). Effective e-learning integration with traditional learning in a blended learning environment. In: E*uropean and Mediterranean Conference on Information Systems,* 1-16.

111. Hammer, F., Egger-Lampl, S., & Möller, S. (2018). Quality-of-user-experience: a position paper. *Quality and User Experience*, 3(1), 1-15.

112. Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives*, 4, 31-52.

113. Hassenzahl, M. (2010). Experience design: Technology for all the right reasons. *Synthesis lectures on human-centered informatics*, 3(1), 1-95.

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

115. Hausman, J. A., Sidak, J. G., & Singer, H. J. (2001). Residential demand for broadband telecommunications and consumer access to unaffiliated Internet content providers. *Yale Journal on Regulation*, 18, 129-173

116. Haybron, D. M. (2008). *The Pursuit of Unhappiness: The Elusive Psychology of Well-being*. Oxford University Press on Demand.

117. Heddaya, A. S. (2002). An Economically Scalable Internet. *Computer*, 35(9), 93-95.

118. Hei, X., Liu, Y., & Ross, K. W. (2008). IPTV over P2P streaming networks: the mesh-pull approach. *IEEE Communications Magazine*, 46(2), 86-92.

119. Hekstra, A. P., Beerends, J. G., Ledermann, D., De Caluwe, F. E., Kohler, S., Koenen, R. H., ... & Schlauss, D. (2002). PVQM–A Perceptual Video Quality Measure. *Signal Processing: Image Communication*, 17(10), 781-798.

120. Hestnes, B., Brooks, P., Heiestad, S., Ulseth, T., & Aaby, C. (2003). Quality of experience in real-time person-person communication–user based QoS expressed in technical network QoS terms. In: *19th International Symposium on Human Factors in Telecommunication* (HFT03), 1-4.

121. Hewage, C. T. & Martini, M. G. (2020). Time varying quality estimation for HTTP based adaptive video streaming. In: *IEEE International Conference on Multimedia & Expo Workshops* (ICMEW), 1-6.

122. Hilgard, E. R. (1980). The trilogy of mind: Cognition, affection, and conation. *Journal of the History of the Behavioral Sciences*, 16(2), 107-117.

123. Hines, A., Skoglund, J., Kokaram, A. C., & Harte, N. (2015). ViSQOL: an objective speech quality model. *EURASIP Journal on Audio, Speech, and Music Processing*, 2015(1), 1-18.

124. Hossain, E. & Hasan, M. (2015). 5G cellular: Key enabling technologies and research challenges. *IEEE Instrumentation & Measurement Magazine*, 18(3), 11-21.

125. Hoßfeld, T., Egger, S., Schatz, R., Fiedler, M., Masuch, K., & Lorentzen, C. (2012). Initial delay vs. interruptions: Between the devil and the deep blue sea. In: *Fourth International Workshop on Quality of Multimedia Experience*, 1-6.

126. Hoßfeld, T., Heegaard, P. E., Varela, M., & Möller, S. (2016). QoE beyond the MOS: an in-depth look at QoE via better metrics and their relation to MOS. *Quality and User Experience*, 1(1), 1-23.

127. Hoßfeld, T., Keimel, C., Hirth, M., Gardlo, B., Habigt, J., Diepold, K., & Tran-Gia, P. (2013). Best practices for QoE crowdtesting: QoE assessment with crowdsourcing. *IEEE Transactions on Multimedia*, 16(2), 541-558.

128. Hoßfeld, T., Schatz, R., & Egger, S. (2011). SOS: The MOS is not enough! In: *Third International Workshop on Quality of Multimedia Experience*, 131-136.

129. Hoßfeld, T., Seufert, M., Sieber, C., & Zinner, T. (2014). Assessing effect sizes of influence factors towards a QoE model for HTTP adaptive streaming. In: *Sixth International Workshop on Quality of Multimedia Experience* (QoMEX), 111-116.

130. Hou, Y. T., Pan, J., Li, B., Tang, X., & Panwar, S. (2002). Modeling and analysis of an expiration-based hierarchical caching system. In: *Global Telecommunications Conference, GLOBECOM'02*, 2468-2472.

131. Huang, J. & Hsu, C. H. (2010). The impact of customer-to-customer interaction on cruise experience and vacation satisfaction. *Journal of Travel Research*, 49(1), 79-92.

132. Ickin, S., Wac, K., Fiedler, M., Janowski, L., Hong, J. H., & Dey, A. K. (2012). Factors influencing quality of experience of commonly used mobile applications. *IEEE Communications Magazine*, 50(4), 48-56.

133. ITU-R (2001). Recommendation BS.1387 (11/2001), Perceptual evaluation of audio quality (PEAQ). International Telecommunication Union, Geneva, Switzerland.

134. ITU-R (2012). Recommendation BT.500-13 (01/2012), Methodology for the subjective assessment of the quality of television pictures. International Telecommunication Union, Geneva, Switzerland.

135. ITU-T (1984). Recommendation P.82 (Extract from the *Blue Book*), Method for Evaluation of Service from the Standpoint of Speech Transmission Quality. International Telecommunication Union, Geneva, Switzerland.

136. ITU-T (1996). Recommendation P.800 (08/96), Methods for subjective determination of transmission quality. International Telecommunication Union, Geneva, Switzerland.

137. ITU-T (2001). Recommendation G.1000 (11/2001), Communications quality of service: A framework and definitions. International Telecommunication Union, Geneva, Switzerland.

138. ITU-T (2001). Recommendation J.144 (03/2001), Objective Perceptual Video Quality Measurement Techniques for Digital Cable Television in the Presence of a Full Reference. International Telecommunication Union, Geneva, Switzerland.

139. ITU-T (2001). Recommendation P.862 (02/2001), Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs. International Telecommunication Union, Geneva, Switzerland.

140. ITU-T (2005). Recommendation J.241 (04/2005), Quality of service ranking and measurement methods for digital video services delivered over broadband IP networks. International Telecommunication Union, Geneva, Switzerland.

141. ITU-T (2008). Recommendation J.247 (08/2008), Objective Perceptual Multimedia Video Quality Measurement in the Presence of a Full Reference. International Telecommunication Union, Geneva, Switzerland.

142. ITU-T (2008). Recommendation E.800 (09/2008), Quality of telecommunication services: concepts, models, objectives and dependability planning – Terms and definitions related to the quality of telecommunication services. International Telecommunication Union, Geneva, Switzerland.

143. ITU-T (2008). Recommendation P.10/G.100, Amendment 2 (07/2008), Vocabulary for performance and quality of service. International Telecommunication Union, Geneva, Switzerland.

144. ITU-T (2011). Recommendation P.863 (01/2011), Perceptual Objective Listening Quality Assessment. International Telecommunication Union, Geneva, Switzerland.

145. ITU-T (2017). Recommendation P.10/G.100 (11/2017), Vocabulary for performance, quality of service and quality of experience. International Telecommunication Union, Geneva, Switzerland.

146. ITU-T (2018). Recommendation P.809 (06/2018), Methods for objective and subjective assessment of speech and video quality, Subjective evaluation methods for gaming quality. International Telecommunication Union, Geneva, Switzerland.

147. Jackson, F. (1982). Epiphenomenal qualia. *The Philosophical Quarterly* (1950-), 32(127), 127-136.

148. Jacoby, J., Szybillo, G. J., & Berning, C. K. (1976). Time and consumer behavior: An interdisciplinary overview. *Journal of Consumer Research*, 2(4), 320-339.

149. Jaeger, G. & Selznick, P. (1964). A normative theory of culture. *American Sociological Review*, 653-669.

150. Jahromi, H. Z., Delaney, D. T., & Hines, A. (2020). Beyond first impressions: Estimating quality of experience for interactive web applications. *IEEE Access*, 8, 47741-47755.

151. Jain, R. (2004). Quality of experience. *IEEE Multimedia*, 11(1), 96-95.

152. Janowski, L. & Papir, Z. (2009). Modeling subjective tests of quality of experience with a generalized linear model. In: *International Workshop on Quality of Multimedia Experience*, 35-40.

153. Jara-Díaz, S. & Rosales-Salas, J. (2017). Beyond transport time: A review of time use modeling. *Transportation Research Part A: Policy and Practice*, 97, 209-230.

154. Jay, M. (2006). *Songs of Experience: Modern American and European Variations on a Universal Theme*. University of California Press.

155. Jiang, J., Sekar, V., & Zhang, H. (2014). Improving fairness, efficiency, and stability in http-based adaptive video streaming with FESTIVE. *IEEE/ACM Transactions on Networking*, 22(1), 326-40.

156. Jin, N., Lee, S., & Lee, H. (2015). The effect of experience quality on perceived value, satisfaction, image and behavioral intention of water park patrons: New versus repeat visitors. *International Journal of Tourism Research*, 17(1), 82-95.

157. Jones, B. L. & McManus, P. R. (1986). Graphic scaling of qualitative terms. *SMPTE Journal*, 95(11), 1166-1171.

158. Juluri, P., Tamarapalli, V., & Medhi, D. (2015). Measurement of quality of experience of video-on-demand services: A survey. *IEEE Communications Surveys & Tutorials*, 18(1), 401-418.

159. Kahneman, D. (2011). Experienced Utility and Objective Happiness: A Moment-Based Approach. Kahneman, D. (2011). In: Kahneman, D. & Tversky, A. (Eds.), *Choices, values, and frames*. New York: Cambridge University Press, 673-692.

160. Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan.

161. Kahneman, D. (1999). Objective happiness. In: Kahneman et al. (Eds.), *Well-Being: The Foundations of Hedonic Psychology*, New York: Russell Sage Foundation, 3-25.

162. Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The Day Reconstruction Method. *Science*, 306(5702), 1776-1780.

163. Kahneman, D. & Snell, J. (1992). Predicting a changing taste: Do people know what they will like? *Journal of Behavioral Decision Making*, 5(3), 187-200.

164. Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, 47(2), 263-92.

165. Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics*, 112(2), 375-406.

166. Kandadai, S., Hardin, J., & Creusere, C. D. (2008). Audio quality assessment using the mean structural similarity measure. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*, 221-224.

167. Katz, J. (1992). Psychophysiological contributions to phantom limbs. *The Canadian Journal of Psychiatry*, 37(5), 282-298.

168. 1Keighrey, C., Flynn, R., Murray, S., & Murray, N. (2017). A QoE evaluation of immersive augmented and virtual reality speech & language assessment applications. In: *Ninth International Conference on Quality of Multimedia Experience* (QoMEX), 1-6.

169. Kekolahti, P., Kilkki, K., Hämmäinen, H., & Riikonen, A. (2016). Features as predictors of phone popularity: an analysis of trends and structural breaks. *Telematics and Informatics*, 33(4), 973-989.

170. Kerpez, K., Waring, D., Lapiotis, G., Lyles, J. B., & Vaidyanathan, R. (2006). IPTV service assurance. *IEEE Communications Magazine*, 44(9), 166-172.

171. Khirman, S. & Henriksen, P. (2002). Relationship between quality-of-service and quality-of-experience for public internet service. In: *Proc. of the 3rd Workshop on Passive and Active Measurement*, 1-6.

172. Kilkki, K. (2007). A practical model for analyzing long tails. *First Monday*.

173. Kilkki, K. (2008). Quality of experience in communications ecosystem. *Journal of Universal Computer Science*, 14(5), 615-624.

174. Kilkki, K. & Finley, B. (2019). In Search of Lost QoS, *arXiv preprint*, arXiv:190106867.

175. Kilkki, K. & Hämmäinen, H. (2019). Value of Time in the Context of Communication Services. In: *European Conference of the International Telecommunication Society (ITS)*, Helsinki, Finland.

176. Kim, H. J. & Choi, S. G. (2010). A study on a QoS/QoE correlation model for QoE evaluation on IPTV service. In: *The 12th International Conference on Advanced Communication Technology* (ICACT), 1377-1382.

177. Kivela, J. & Crotts, J. C. (2006). Tourism and gastronomy: Gastronomy's influence on how tourists experience a destination. *Journal of Hospitality & Tourism Research*, 30(3), 354-377.

178. Konradt, U., Filip, R., & Hoffmann, S. (2003). Flow experience and positive affect during hypermedia learning. *British Journal of Educational Technology*, 34(3), 309-327.

179. Kraut, R. (2009). *What is good and why*. Harvard University Press.

180. Kuipers, F., Kooij, R., Vleeschauwer, D. D., & Brunnström, K. (2010). Techniques for measuring quality of experience. In: *International Conference on Wired/Wireless Internet Communications*, 216-227.

181. Laghari, K. U. R. & Connelly, K. (2012). Toward total quality of experience: A QoE model in a communication ecosystem. *IEEE Communications Magazine*, 50(4), 58-65.

182. Lallemand, C., Gronier, G., & Koenig, V. (2015). User experience: A concept without consensus? Exploring practitioners' perspectives through an international survey. *Computers in Human Behavior*, 43, 35-48.

183. Larson, R., Csikszentmihalyi, M. (2014). The experience sampling method. In: Reis, H. (Ed.), *Naturalistic Approaches to Studying Social Interaction*, 41-56.

184. Lee, Y., Kozar, K. A., & Larsen, K. R. (2003). The technology acceptance model: Past, present, and future. *Communications of the Association for Information Systems*, 12(1), 50.

185. Lee, Y. T., Chen, K. T., Su, H. I., & Lei, C. L. (2012). Are all games equally cloud-gaming-friendly? An electromyographic approach. In: *11th Annual Workshop on Network and Systems Support for Games* (NetGames), 1-6.

186. Levine, R. N. (2008). *A geography of time: On tempo, culture, and the pace of life*. Basic Books.

187. Li, Z., Aaron, A., Katsavounidis, I., Moorthy, A., & Manohara, M. (2016). Toward a practical perceptual video quality metric. *The Netflix Tech Blog*, 6(2).

188. Li-yuan, L., Wen-an, Z., & Jun-de, S. (2006). The research of quality of experience evaluation method in pervasive computing environment. In: *First International Symposium on Pervasive Computing and Applications*, 178-182.

189. Ljubojevic, M., Vaskovic, V., Stankovic, S., & Vaskovic, J. (2014). Using supplementary video in multimedia instruction as a teaching tool to increase efficiency of learning and quality of experience. *International Review of Research in Open and Distributed Learning*, 15(3), 275-291.

190. Lu, Z., Lin, W., Ong, E., Yang, X., & Yao, S. (2003). PQSM-based RR and NR video quality metrics. In: *Visual Communications and Image Processing 2003*, 633-640.

191. Lumsdon, L. M. & McGrath, P. (2011). Developing a conceptual framework for slow travel: A grounded theory approach. *Journal of Sustainable Tourism*, 19(3), 265-279.

192. Luo, K., Dang, S., Shihada, B., & Alouini, M. S. (2019). Prospect Theory for Human-Centric Communications. *arXiv preprint*, arXiv:191207219.

193. Mahmud, R., Kotagiri, R., & Buyya, R. (2018). Fog computing: A taxonomy, survey and future directions. In: *Internet of everything*, Springer, Singapore, 103-130.

194. Maia, O. B., Yehia, H. C., & de Errico, L. (2015). A concise review of the quality of experience assessment for video streaming. *Computer Communications*, 57, 1-12.

195. Mano, H. (1999). The influence of pre-existing negative affect on store purchase intentions. *Journal of Retailing*, 75(2), 149-172.

196. Mao, H., Netravali, R., & Alizadeh, M. (2017). Neural adaptive video streaming with pensieve. In: *Proceedings of the Conference of the ACM Special Interest Group on Data Communication*, 197-210.

197. Martela, F. & Steger, M. F. (2016). The three meanings of meaning in life: Distinguishing coherence, purpose, and significance. *The Journal of Positive Psychology*, 11(5), 531-545.

198. Martens, H. & Martens, M. (2001). *Multivariate analysis of quality*. Wiley, Chichester.

199. Mathwick, C., Malhotra, N., & Rigdon, E. (2001). Experiential value: conceptualization, measurement and application in the catalog and Internet shopping environment. *Journal of Retailing*, 77(1), 39-56.

200. Mazzucato, M. (2018). *The value of everything: Making and taking in the global economy*, Hachette UK.

201. McCarthy, J. & Wright, P. (2004). *Technology as Experience*. Cambridge, MA: MIT Press.

202. McCurdy, H. G. (1950). Consciousness and the galvanometer. *Psychological Review*, 57(6), 322-327.

203. McNamara, N. & Kirakowski, J. (2005). Defining usability: quality of use or quality of experience? In: *Proceedings of International Professional Communication Conference*, 200-204.

204. Melzack, R. (1993). Pain: Past, Present and Future. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 47(4), 615-29.

205. Meyer, D. K. & Turner, J. C. (2002). Discovering emotion in classroom motivation research. *Educational psychologist*, 37(2), 107-114.

206. Min, X., Zhai, G., Zhou, J., Farias, M. C., & Bovik, A. C. (2020). Study of subjective and objective quality assessment of audio-visual signals. *IEEE Transactions on Image Processing*, 29, 6054-6068.

207. Minhas, T. N. & Fiedler, M. (2013). Quality of experience hourglass model. In: *International Conference on Computing, Management and Telecommunications (ComManTel)*, 87-92.

208. Mitra, K., Zaslavsky, A., & Åhlund, C. (2013). Context-aware QoE modelling, measurement, and prediction in mobile computing systems. *IEEE Transactions on Mobile Computing*, 14(5), 920-936.

209. Mittal, A., Moorthy, A. K., & Bovik, A. C. (2012). No-reference image quality assessment in the spatial domain. *IEEE Transactions on image processing*, 21(12), 4695-4708.

210. Mittal, A., Soundararajan, R., & Bovik, A. C. (2012). Making a "completely blind" image quality analyzer. *IEEE Signal Processing Letters*, 20(3), 209-212.

211. Mlodinow, L. (2013). *Subliminal: How your unconscious mind rules your behavior*. Vintage.

212. Mok, R. K., Chan, E. W., & Chang, R. K. (2011). Measuring the Quality of Experience of HTTP Video Streaming. In: *12th IFIP/IEEE International Symposium on Integrated Network Management (IM 2011) and Workshops*, 485-492.

213. Mok, R. K., Chan, E. W., Luo, X., & Chang, R. K. (2011). Inferring the QoE of HTTP video streaming from user-viewing activities. In: *Proceedings of the First ACM SIGCOMM Workshop on Measurements up the Stack*, 31-36.

214. Möller, S., Engelbrecht, K. P., Kuhnel, C., Wechsung, I., & Weiss, B. (2009). A taxonomy of quality of service and quality of experience of multimodal human-machine interaction. In: *International Workshop on Quality of Multimedia Experience*, 7-12.

215. Möller, S. & Köster, F. (2017). Review of recent standardization activities in speech quality of experience. *Quality and User Experience*, 2(1), 1-18.

216. Möller, S., Schmidt, S., & Zadtootaghaj, S. (2018). New ITU-T standards for gaming QoE evaluation and management. In: *Tenth International Conference on Quality of Multimedia Experience* (QoMEX), 1-6.

217. Moon, S. E. & Lee, J. S. (2015). Perceptual experience analysis for tone-mapped HDR videos based on EEG and peripheral physiological signals. *IEEE Transactions on Autonomous Mental Development*, 7(3), 236-247.

218. Moorthy, A. K. & Bovik, A. C. (2011). Blind image quality assessment: From natural scene statistics to perceptual quality. *IEEE Transactions on Image Processing*, 20(12), 3350-3364.

219. 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. *International Journal of Human-Computer Studies*, 54(6), 877-901.

220. Morrissey, K., McCarthy, J., & Pantidi, N. (2017). The value of experience-centred design approaches in dementia research contexts. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 1326-1338.

221. Mushtaq, M. S., Augustin, B., & Mellouk, A. (2012). Empirical study based on machine learning approach to assess the QoS/QoE correlation. In: *17th European Conference on Networks and Optical Communications*, 1-7.

222. Nahrstedt, K. & Smith J. M. (1995). The QoS Broker. *IEEE Multimedia,* 2(1), 53-67.

223. Nakamura, J. & Csikszentmihalyi, M. (2002). The concept of flow. In: Snyder C. R. & Lopez S. J. (Eds.), *Handbook of Positive Psychology*, New York: Oxford University Press, 89-105.

224. Nawaz, O., Fiedler, M., De Moor, K., & Khatibi, S. (2020). Influence of Gender and Viewing Frequency on Quality of Experience. In: *Twelfth International Conference on Quality of Multimedia Experience* (QoMEX), 1-4.

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

226. O'Neil, T. M. (2002). Quality of Experience and Quality of Service for IP Video Conferencing. Whitepaper, Polycom.

227. Orsolic, I., Pevec, D., Suznjevic, M., & Skorin-Kapov, L. (2017). A machine learning approach to classifying YouTube QoE based on encrypted network traffic. *Multimedia Tools and Applications*, 76(21), 22267-22301.

228. Otto, J. E. & Ritchie, J. B. (1996). The service experience in tourism. *Tourism Management*, 17(3), 165-174.

229. Oyman, O. & Singh, S. (2012). Quality of Experience for HTTP Adaptive Streaming Services. *IEEE Communications Magazine*, 50(4), 20-27.

230. Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A conceptual model of service quality and its implications for future research. *Journal of Marketing*, 49(4), 41-50.

231. Parfit, D. (2011). *On What Matters* (Vol. 1). Oxford University Press.

232. Pauliks, R., Tretjaks, K., Belahs, K., & Pauliks, R. (2013). A survey on some measurement methods for subjective video quality assessment. In: *World Congress on Computer and Information Technology* (WCCIT), 1-6.

233. Perkis, A. et al. (2020). QUALINET white paper on definitions of immersive media experience (IMEx). *arXiv preprint*, arXiv:2007.07032.

234. Phan, T. K., Griffin, D., Maini, E., & Rio, M. (2017). Utility-Centric Networking: Balancing Transit Costs with Quality of Experience. *IEEE/ACM Transactions on Networking*, 26(1), 245-258.

235. Pilling, D. (2019). *The growth delusion: The wealth and well-being of nations*. Bloomsbury Publishing.

236. Pilloni, V., Floris, A., Meloni, A., & Atzori, L. (2018). Smart Home Energy Management Including Renewable Sources: A QoE-Driven Approach. *IEEE Transactions on Smart Grid*, 9(3), 2006-2018.

237. Pine, B. J. & Gilmore G. H. (1998). Welcome to the experience economy. *Harvard Business Review*, 76, 97-105.

238. Pinson, M. H. & Wolf, S. (2003). An objective method for combining multiple subjective data sets. In: *Visual Communications and Image Processing* 2003, 583-592.

239. Pinson, M. H. & Wolf, S. (2004). A new standardized method for objectively measuring video quality. *IEEE Transactions on broadcasting*, 50(3), 312-322.

240. Pohjola, O. P. & Kilkki, K. (2007). Value-based methodology to analyze communication services. *NETNOMICS: Economic Research and Electronic Networking*, 8(1), 135-151.

241. Porcu, S., Floris, A., Voigt-Antons, J. N., Atzori, L., & Möller, S. (2020). Estimation of the Quality of Experience during video streaming from facial expression and gaze direction. *IEEE Transactions on Network and Service Management*, 17(4), 2702-2716.

242. Porter, M. E. (2008). *On Competition*. Harvard Business Press.

243. Prinz, J. (2012). *The Conscious Brain: How Attention Engenders Experience*, Oxford University Press.

244. Raake, A. & Egger, S. (2014). Quality and Quality of Experience. In: Möller, S. & Raake, A. (Eds.), *Quality of Experience: Advanced Concepts, Applications and Methods*, Springer, Cham, 11-33.

245. Raheel, A., Majid, M., Alnowami, M., & Anwar, S. M. (2020). Physiological sensors based emotion recognition while experiencing tactile enhanced multimedia. *Sensors*, 20(14), 4037.

246. Razaak, M., Martini, M. G., & Savino, K. (2014). A study on quality assessment for medical ultrasound video compressed via HEVC. *IEEE Journal of Biomedical and Health Informatics*, 18(5), 1552-1559.

247. Read, D. (2007). Experienced utility: utility theory from Jeremy Bentham to Daniel Kahneman. *Thinking & Reasoning*, 13(1), 45-61.

248. Reeves, C. A. & Bednar, D. A. (1994). Defining quality: alternatives and implications. *Academy of Management Review*, 19(3), 419-445.

249. Reichl, P., Egger, S., Schatz, R., & D'Alconzo, A. (2010). The logarithmic nature of QoE and the role of the Weber-Fechner law in QoE assessment. In: *IEEE International Conference on Communications*, 1-5.

250. Reichl, P., Tuffin, B., & Schatz, R. (2013). Logarithmic laws in service quality perception: where microeconomics meets psychophysics and quality of experience. *Telecommunication Systems*, 52(2), 587-600.

251. Reiter, U., Brunnström, K., De Moor, K., Larabi, M. C., Pereira, M., Pinheiro, A., & Zgank, A. (2014). Factors influencing quality of experience. In: Möller, S. & Raake, A. (Eds.), *Quality of Experience: Advanced Concepts, Applications and Methods*, Springer, Cham, 55-72.

252. Riva, G., Baños, R. M., Botella, C., Wiederhold, B. K., & Gaggioli, A. (2012). Positive technology: using interactive technologies to promote positive functioning. *Cyberpsychology, Behavior, and Social Networking*, 15(2), 69-77.

253. Rubino, G. (2006). Quantifying the quality of audio and video transmissions over the Internet: the PSQA approach. *Communication Networks and Computer Systems*: A Tribute to Professor Erol Gelenbe, 235-250.

254. Ryan, R. M. & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54-67.

255. Saad, M. A., Bovik, A. C., & Charrier, C. (2012). Blind image quality assessment: A natural scene statistics approach in the DCT domain. *IEEE Transactions on Image Processing*, 21(8), 3339-3352.

256. Sabet, S. S., Schmidt, S., Zadtootaghaj, S., Naderi, B., Griwodz, C., & Möller, S. (2020). A latency compensation technique based on game characteristics to mitigate the influence of delay on cloud gaming quality of experience. In: *Proceedings of the 11th ACM Multimedia Systems Conference*, 15-25.

257. Sackl, A. & Schatz, R. (2013). Evaluating the impact of expectations on end-user quality perception. In: *Proceedings of Forth International Workshop on Perceptual Quality of Systems (PQS)*, 122-128.

258. Sackl, A., Schatz, R., & Raake, A. (2017). More than I ever wanted or just good enough? User expectations and subjective quality perception in the context of networked multimedia services. *Quality and User Experience*, 2(1), 1-27.

259. Sahai, A., Ouyang, J., Machiraju, V., & Wurster, K. (2001). Specifying and guaranteeing quality of service for web services through real time measurement and adaptive control. *Hewlett-Packard Labs Technical Report.*

260. Salgado, D. P., Martins, F. R., Rodrigues, T. B., Keighrey, C., Flynn, R., Naves, E. L. M., & Murray, N. (2018). A QoE assessment method based on EDA, heart rate and EEG of a virtual reality assistive technology system. In: *Proceedings of the 9th ACM Multimedia Systems Conference*, 517-520.

261. Schatz, R., Egger, S., & Platzer, A. (2011). Poor, good enough or even better? Bridging the gap between acceptability and QoE of mobile broadband data services. In: *IEEE International Conference on Communications (ICC)*, 1-6.

262. Schatz, R., Hoßfeld, T., Janowski, L., & Egger, S. (2013). From packets to people: Quality of Experience as a new measurement challenge. In: *Data Traffic Monitoring and Analysis*, 219-263.

263. Schiefele, U. (1991). Interest, learning, and motivation. *Educational Psychologist*, 26(3-4), 299-323.

264. Schmitt, B. (1999). Experiential marketing. *Journal of Marketing Management*, 15(1-3), 53-67.

265. Seshadrinathan, K. & Bovik, A. C. (2009). Motion tuned spatio-temporal quality assessment of natural videos. *IEEE Transactions on Image Processing*, 19(2), 335-350.

266. Seufert, M., Egger, S., Slanina, M., Zinner, T., Hoßfeld, T., & Tran-Gia, P. (2014). A Survey on Quality of Experience of HTTP Adaptive Streaming. *IEEE Communications Surveys & Tutorials*, 17(1), 469-492.

267. Seufert, M. (2019). Fundamental advantages of considering quality of experience distributions over mean opinion scores. In: *Eleventh International Conference on Quality of Multimedia Experience* (QoMEX), 1-6.

268. Seufert, M. (2021). Statistical methods and models based on quality of experience distributions. *Quality and User Experience*, 6(1), 1-27.

269. Shackel, B. (1991). Usability - Context, Framework, Definition, Design and Evaluation. In: B Shackel, B. & Richardson, S. (Eds.), *Human Factors for Informatics Usability*, 21-37.

270. Shah-Mansouri, H. & Wong, V. W. (2018). Hierarchical fog-cloud computing for IoT systems: A computation offloading game. *IEEE Internet of Things Journal*, 5(4), 3246-3257.

271. Shao, F., Lin, W., Gu, S., Jiang, G., & Srikanthan, T. (2013). Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics. *IEEE Transactions on Image Processing*, 22(5), 1940-1953.

272. Sheikh, H. R. & Bovik, A. C. (2006). Image information and visual quality. *IEEE Transactions on Image Processing*, 15(2), 430-444.

273. Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly*, 18, 158-176.

274. Shin, D. (2017). Conceptualizing and measuring Quality of Experience of the Internet of Things: Exploring how quality is perceived by users. *Information & Management*, 54(8), 998-1011.

275. Shin, D. (2018). Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? *Computers in Human Behavior*, 78, 64-73.

276. Shin, D. (2019). How does immersion work in augmented reality games? A user-centric view of immersion and engagement. *Information, Communication & Society*, 22(9), 1212-1229.

277. Shuttenberg, E. M. & Poppenhagen, B. W. (1980). Current theory and research in experiential learning for adults. *Journal of Experiential Education*, 3(1), 27-31.

278. Siegel, D. J. (2007). *The mindful brain: Reflection and attunement in the cultivation of well-being*. WW Norton & Company.

279. Siller, M. & Woods, J. (2003). Improving Quality of Experience for multimedia services by QoS arbitration on a QoE framework. In: *Proceedings of 13th Packed Video Workshop*, Nantes, France.

280. Simmel, G. (1978). *The Philosophy of Money*. Routledge.

281. Skinner, B. F. (1976). *About Behaviorism*. Vintage Books.

282. Skorin-Kapov, L., Varela, M., Hoßfeld, T., & Chen, K. T. (2018). A survey of emerging concepts and challenges for QoE management of multimedia services. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 14(2s), 1-29.

283. Small, K. A. (2012). Valuation of travel time. *Economics of Transportation*, 1(1-2), 2-14.

284. Song, W., Tjondronegoro, D. W., & Docherty, M. (2012). Understanding user experience of mobile video: framework, measurement, and optimization. *Mobile Multimedia: User and Technology Perspectives*, 3-30.

285. Soundararajan, R. & Bovik, A. C. (2012). RRED indices: Reduced reference entropic differencing for image quality assessment. *IEEE Transactions on Image Processing*, 21(2), 517-526.

286. Soundararajan, R. & Bovik, A. C. (2013). Video quality assessment by reduced reference spatio-temporal entropic differencing. *IEEE Transactions on Circuits and Systems for Video Technology*, 23(4), 684-694.

287. Stacey, R. D. (1996). *Complexity and creativity in organizations*. Berrett-Koehler Publishers.

288. Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103(2684), 677-680.

289. Stigler, G. J. (1950). The development of utility theory. I. *Journal of Political Economy*, 58(4), 307-327.

290. Streijl, R. C., Winkler, S., & Hands, D. S. (2016). Mean Opinion Score (MOS) revisited: methods and applications, limitations and alternatives. *Multimedia Systems*, 22(2), 213-227.

291. Sullivan, M., Pratt, J., & Kortum, P. (2008). Practical issues in subjective video quality evaluation: Human factors vs. psychophysical image quality evaluation. In: *Proceedings of the 1st International Conference on Designing Interactive User Experiences for TV and Video*, 1-4.

292. Sumner, L. W. (1996). *Welfare, Happiness, and Ethics*. Clarendon Press.

293. Suri, J. F. (2003). The Experience of Evolution: Developments in Design Practice. *The Design Journal*, 6(2), 39-48.

294. Takahashi, A., Hands, D., & Barriac, V. (2008). Standardization activities in the ITU for a QoE assessment of IPTV. *IEEE Communications Magazine*, 46(2), 78-84.

295. Takatalo, J., Häkkinen, J., Kaistinen, J., & Nyman, G. (2007). Measuring user experience in digital gaming: Theoretical and methodological issues. In: *Image Quality and System Performance IV* (6494), 2-13.

296. Talbott, S. (2007). *Devices of the Soul: Battling for Our Selves in an Age of Machines*. O'Reilly Media.

297. Thakolsri, S., Kellerer, W., & Steinbach, E. (2011). QoE-based cross-layer optimization of wireless video with unperceivable temporal video quality fluctuation. In: *IEEE International Conference on Communications (ICC)*, 1-6.

298. Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science*, 4(3), 199-214.

299. Thanou, A., Tsiropoulou E. E., & Papavassiliou, S. (2019). Quality of experience under a prospect theoretic perspective: A cultural heritage space use case. *IEEE Transaction on Computational Social Systems*, (61), 135-148.

300. Thompson, E. (2010). *Mind in Life*. Cambridge, MA: Harvard University Press.

301. Thurstone, L. L. (1927). A law of comparative judgment. *Psychological review*, 34(4), 266-270.

302. Tian-Cole, S., Crompton, J. L., & Willson, V. L. (2002). An empirical investigation of the relationships between service quality, satisfaction and behavioral intentions among visitors to a wildlife refuge. *Journal of Leisure Research*, 34(1), 1-24.

303. Tian-Cole, S. & Chancellor, H. C. (2009). Examining the festival attributes that impact visitor experience, satisfaction and re-visit intention. *Journal of Vacation Marketing*, 15(4), 323-333.

304. Tononi, G. & Koch, C. (2015). Consciousness: here, there and everywhere? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1668), 20140167.

305. Torbert, W. (1974). Doing Rawls justice. *Harvard Educational Review*, 44(4), 459-469.

306. Tsiropoulou, E. E., Kousis, G., Thanou, A., Lykourentzou, I., & Papavassiliou, S. (2018). Quality of Experience in cyber-physical social systems based on reinforcement learning and game theory. *Future Internet*, 10(11), 108.

307. Tsolkas, D., Liotou, E., Passas, N., & Merakos, L. (2017). A survey on parametric QoE estimation for popular services. *Journal of Network and Computer Applications*, 77, 1-17.

308. Tversky, A. (2018). *The Essential Tversky*. MIT Press.

309. Um, S., Chon, K. & Ro, Y. (2006). Antecedents of revisit intention. *Annals of Tourism Research*, 33(4), 1141-1158.

310. van den Branden Lambrecht, C. J. (1996). Perceptual models and architectures for video coding applications. Ph.D. Thesis, Swiss Federal Institute of Technology, Lausanne.

311. van den Branden Lambrecht, C. J., & Verscheure, O. (1996). Perceptual quality measure using a spatiotemporal model of the human visual system. In: *Digital Video Compression: Algorithms and Technologies* (Vol. 2668), 450-461.

312. van Moorsel, A. (2001). Metrics for the Internet age: Quality of Experience and Quality of Business. In: *5th Performability Workshop*, Erlagen, Germany.

313. Varela, M., Skorin-Kapov, L., De Moor, K., & Reichl, P. (2016). QoE—Defining a user-centric concept for service quality. In: Chen, C. W. et al. (Eds.), *Multimedia Quality of Experience (QoE): Current Status and Future Requirements*, John Wiley & Sons, 5-28.

314. Vasileva-Stojanovska, T., Vasileva, M., Malinovski, T., & Trajkovik, V. (2015). An ANFIS model of quality of experience prediction in education. *Applied Soft Computing*, 34, 129-138.

315. Videnovik, M., Trajkovik, V., Kiønig, L. V., & Vold, T. (2020). Increasing quality of learning experience using augmented reality educational games. *Multimedia Tools and Applications*, 79(33), 23861-23885.

316. Video Quality Experts Group (VQEG) (2000). Final report from the video quality experts group on the validation of objective models of video quality assessment.

317. Voran, S. (1991). The development of objective video quality measures that emulate human perception. In: *IEEE GLOBECOM*, 1776-1781.

318. Walls, A. R., Okumus, F., Wang, Y. R., & Kwun, D. J. W. (2011). An epistemological view of consumer experiences. *International Journal of Hospitality Management*, 30(1), 10-21.

319. Waltl, M., Rainer, B., Timmerer, C., & Hellwagner, H. (2013). An end-to-end tool chain for Sensory Experience based on MPEG-V. *Signal Processing: Image Communication*, 28(2), 136-150.

320. Wang, K., Barkowsky, M., Brunnström, K., Sjöström, M., Cousseau, R., & Le Callet, P. (2012). Perceived 3D TV transmission quality assessment: multi-laboratory results using absolute category rating on quality of experience scale. *IEEE Transactions on Broadcasting*, 58(4), 544-557.

321. Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600-612.

322. Wang, Z., Lu, L., & Bovik, A. C. (2004). Video quality assessment based on structural distortion measurement. Signal processing: *Image Communication*, 19(2), 121-132.

323. Wang, Z., Simoncelli, E. P., & Bovik, A. C. (2003). Multiscale structural similarity for image quality assessment. In: *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers*, 1398-1402.

324. Watson, A. & Sasse, M. A. (1998). Measuring perceived quality of speech and video in multimedia conferencing applications. In: *Proceedings of the Sixth ACM International Conference on Multimedia*, 55-60.

325. Watson, A. B., Hu, Q. J., & McGowan III, J. F. (2001). Digital video quality metric based on human vision. *Journal of Electronic Imaging*, 10(1), 20-29.

326. Wechsung, I. & De Moor, K. (2014). Quality of experience versus user experience. In: Möller, S. & Raake, A. (Eds.), *Quality of Experience: Advanced Concepts, Applications and Methods*, Springer, Cham, 35-54.

327. Whalen, T. E., Noël, S., & Stewart, J. (2003). Measuring the human side of virtual reality. In: *IEEE International Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems* (VECIMS'03), 8-12.

328. Whillans, A. V., Dunn, E. W., Smeets, P., Bekkers, R., & Norton, M. I. (2017). Buying time promotes happiness. In: *Proceedings of the National Academy of Sciences*, 114(32), 8523-8527.

329. Winkler, S. (1998). A perceptual distortion metric for digital color images. In: *Proceedings 1998 International Conference on Image Processing* (ICIP98), 399-403.

330. Winkler, S. & Mohandas, P. (2008). The evolution of video quality measurement: From PSNR to hybrid metrics. *IEEE Transactions on Broadcasting*, 54(3), 660-668.

331. Wolter, K. & van Moorsel, A. (2001). The relationship between Quality of Service and Business Metrics: Monitoring, Notification and Optimization. *Hewlett-Packard Labs Technical Report.*

332. Wu, H. C., Li, M. Y., & Li, T. (2018). A study of experiential quality, experiential value, experiential satisfaction, theme park image, and revisit intention. *Journal of Hospitality & Tourism Research*, 42(1), 26-73.

333. Wu, W., Arefin, A., Rivas, R., Nahrstedt, K., Sheppard, R., & Yang, Z. (2009). Quality of experience in distributed interactive multimedia environments: toward a theoretical framework. In: *Proceedings of the 17th ACM International Conference on Multimedia*, 481-490.

334. Xiao, X. & Ni, L. M. (1999). Internet QoS: A Big Picture. *IEEE Network*, 13(2), 8-18.

335. Xie, J., Yu, F. R., Huang, T., Xie, R., Liu, J., Wang, C., & Liu, Y. (2018). A survey of machine learning techniques applied to software defined networking (SDN): Research issues and challenges. *IEEE Communications Surveys & Tutorials*, 21(1), 393-430.

336. Xu, C., Quan, W., Zhang, H., & Grieco, L. A. (2016). GrIMS: Green information-centric multimedia streaming framework in vehicular ad hoc networks. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(2), 483-498.

337. Xu, J., Xing, L., Perkis, A., & Jiang, Y. (2011). On the properties of mean opinion scores for Quality of Experience management. In: *IEEE International Symposium on Multimedia*, 500-505.

338. Xu, J. B. & Chan, A. (2010). A conceptual framework of hotel experience and customer-based brand equity: Some research questions and implications. *International Journal of Contemporary Hospitality Management*, 22(2), 174-93.

339. Xue, W., Zhang, L., Mou, X., & Bovik, A. C. (2013). Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE Transactions on Image Processing*, 23(2), 684-695.

340. Yin, X., Jindal, A., Sekar, V., & Sinopoli, B. (2015). A control-theoretic approach for dynamic adaptive video streaming over HTTP. In: *Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication*, 325-338.

341. You, J., Reiter, U., Hannuksela, M. M., Gabbouj, M., & Perkis, A. (2010). Perceptual-based quality assessment for audio–visual services: A survey. *Signal Processing: Image Communication*, 25(7), 482-501.

342. Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1993). The nature and determinants of customer expectations of service. *Journal of the Academy of Marketing Science*, 21(1), 1-12.

343. Zerman, E., Valenzise, G., & Dufaux, F. (2017). An extensive performance evaluation of full-reference HDR image quality metrics. *Quality and User Experience*, 2(1), 1-16.

344. Zhang, C., Hoel, A. S., & Perkis, A. (2016). Experiential Qualities and Quality of Experience in Story-telling, and Their Measurability. In: *Proceedings of 5th ISCA/DEGA Workshop on Perceptual Quality of Systems* (PQS 2016), 49-53.

345. Zhang, C., Hoel, A. S., & Perkis, A. (2017). Quality of alternate reality experience and its QoE influencing factors. In: P*roceedings of the 2nd International Workshop on Multimedia Alternate Realities*, 3-8.

346. Zhang, H., Dong, L., Gao, G., Hu, H., Wen, Y., & Guan, K. (2020). DeepQoE: A multimodal learning framework for video quality of experience (QoE) prediction. *IEEE Transactions on Multimedia*, 22(12), 3210-3223.

347. Zhang, L., Zhang, L., Mou, X., & Zhang, D. (2011). FSIM: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, 20(8), 2378-2386.

348. Zhao, T., Liu, Q., & Chen, C. W. (2016). QoE in video transmission: A user experience-driven strategy. *IEEE Communications Surveys & Tutorials*, 19(1), 285-302.

349. Zhao, Z., Cumino, P., Souza, A., Rosario, D., Braun, T., Cerqueira, E., & Gerla, M. (2019). Software-defined unmanned aerial vehicles networking for video dissemination services. *Ad Hoc Networks*, 83, 68-77.

350. Zheleva, A., Durnez, W., Bombeke, K., van Wallendael, G., & De Marez, L. (2020). Seeing is believing: the effect of video quality on quality of experience in virtual reality. In: *Twelfth International Conference on Quality of Multimedia Experience* (QoMEX), 1-4.

351. Zhu, Y., Heynderickx, I., & Redi, J. A. (2015). Understanding the role of social context and user factors in video quality of experience. *Computers in Human Behavior*, 49, 412-426.

352. Zielinski, S., Brooks, P., & Rumsey, F. (2007). On the use of graphic scales in modern listening tests. In: *Audio Engineering Society Convention 123*, Audio Engineering Society.

Abbreviations[75]

ANOVA	analysis of variance
DRM	Day Reconstruction Method
DSCQS	double-stimulus continuous quality-scale
ECG	electrocardiography
EDA	electrodermal activity
EEG	electroencephalography
EMG	electromyography
ESM	Experience Sampling Method
ETSI	European Telecommunications Standards Institute
fMRI	functional magnetic resonance imaging
GSR	galvanic skin response
HAS	HTTP adaptive streaming
HCI	human-computer interaction
HP	Hewlett-Packard
HTTP	Hypertext Transfer Protocol
IETF	Internet Engineering Task Force
IPTV	Internet Protocol television
ITU	International Telecommunication Union
ITU-R	ITU Radiocommunication Sector
ITU-T	ITU Telecommunication Standardization Sector
JND	just-noticeable difference
MOS	mean opinion score
NIRS	near-infrared spectroscopy
NPS	net promoter score
PPG	photoplethysmography
PSNR	peak signal-to-noise ratio
QoBiz	quality of business
QoE	Quality of Experience
QoMEX	International Conferences on Quality of Multimedia Experience
QoP	Quality of Perception
QoS	Quality of Service
QoSE	QoS experienced by the user
RFC	Request for Comments
RMSE	root-mean-square error
SAM	Self-Assessment Manikin
SEM	structural equation modeling
SSIM	Structural Similarity index measure
TAM	technology acceptance model
UX	user experience
VQA	video quality assessment
VQM	video quality metric; video quality measure; video quality model; video quality monitor

[75] In addition, Table A4.2. includes several abbreviations related to the methods to assess the perceived quality of audio and video streams. An abbreviation from Table A4.2 is included in this list only if it appears in other parts of the report (references are excluded).

Appendix 1. Quality of experience in literature

This appendix describes the results of a literature search of the term quality of experience. The search process contained three parts, Cambridge Dictionary of Philosophy, non-fiction books, and quality of experience articles.

The Cambridge Dictionary of Philosophy

The first part of the search was made in R. Audi (general editor), The Cambridge Dictionary of Philosophy, 2nd Edition, 1999. The dictionary consists of more than 4000 entries of concepts and philosophers. Quality of experience is not included as an entry, even experience is lacking. The entry of qualities provides the following argumentation:

- **Qualities**, properties or characteristics (p. 762-763): [76]
  (1) Qualities are physical properties, logical constructions, or dispositions.
  (2) Assuming a representative theory of perception, as Locke did, qualities have two characteristics: qualities of powers (or dispositions) of objects to produce sensory experiences (sensedata on some theories) in humans, and in sensory experience, qualities are represented as intrinsic properties of objects. Intrinsic properties of objects are properties of objects that objects have independently of their environment.

As far as I can assess, none of the entries in the dictionary regarding philosophers mentions the concept of quality of experience. The two concept entries mentioning quality of experience are:

- **Erlebnis** (p. 280):
  As used in post-Hegelian German philosophy, the term [Erlebnis] describes two aspects of subjectivity. The first concerns the epistemology of the human sciences and phenomenology. Against naturalism and objectivism, philosophers appeal to the ineliminable, subjective qualities of experience to argue that interpreters must understand "what it is like to be" some experiencing subject, from the inside.

- **Ethical objectivism** (p. 284):
  The naturalist objectivist believes that values, duties, or whatever are natural phenomena detectable by introspection, perception, or scientific inference. Thus values may be identified with certain empirical qualities of (anybody's) experience, or duties with empirical facts about the effects of action, e.g. as promoting or hindering social cohesion.

Qualia is the closest term for quality of experience with its own article:

- **Qualia** (p. 762):
  Those properties of mental states or events, which determine "what is like" to have them. Sometimes 'phenomenal properties' and 'qualitative features' are used with the same meaning.

In conclusion, quality of experience is not a central concept in philosophy.

Non-fiction books

This part of the appendix is based on my collection of 362 non-fiction books from different disciplines including philosophy, psychology, economics, physiology, sociology, history, biology, technology, and politics. I attempted to find all the cases that contain the phrase quality of experience (also the quality of the experience and the quality of our experience were accepted). I checked all the books in which quality or experience appeared in the index of a book, because "quality of experience" appeared only once as an own entry (in Xiao's book about Quality of Service [334]). Of the 295 books with an index, 10 books have both quality and experience in the index, 45 books have experience but not quality in the index, and 25 books have quality but not experience in the index.

[76] The entry, for some reason, concentrates on the ideas presented by John Locke. The length of both parts, (1) and (2), is about half a page.

I also used the search feature in Google Books (https://books.google.com/) to find cases in which quality of experience is occasionally used, particularly in those books that have experience as an own entry in the index. The results of the search effort are presented in Table A1.1.

Table A1.1: Examples of quality of experience in non-fiction literature.

Author(s)	Title [ref]	Page	Excerpt
Baucells, Manel; Sarin, Rakesh	Engineering Happiness: A New Approach for Building a Joyful Life [34]	17	Economic indicators such as GNP focus largely on market transactions and thus are biased in favor of production and consumption. In contrast, GNH [Gross National Happiness] attempts to measure the quality of human experience and well-being in its totality.
		19	As we have all experienced, good company enhances the quality of any good experience. You enjoy the view from the beach, the concert, or even the cup of coffee more if you share those experiences with the right people.
De Sousa, Ronald	The Rationality of Emotion [74]	260	On the contextualist view advocated by Robert Kraut (unpub.), it is literally impossible for a male to experience female anger or female jealousy, or vice versa, even if his internal states, including the quality of his experience, could somehow be shown to be the same as a jealous or angry woman's.
Dewey, John	Art as Experience [77]	135	In case the term [beauty] is used in theory to designate the total esthetic quality of an experience, it is surely better to deal with the experience itself and show whence and how the quality proceeds. In that case, beauty is the response to that which to reflection is the consummated movement of matter integrated through its inner relations into a single qualitative whole.
		224	But just such feelings, and what other writers have called organic "clicks," are the gross indication of complete organic participation, while it is the fullness and immediacy of this participation that constitutes the esthetic quality of an experience, just as it is that which transcends the intellectual.
		305	I have had occasion to speak more than once of a quality of intense esthetic experience that is so immediate as to be ineffable and mystical. An intellectualized rendering of this immediate quality of experience translates it into the terms of a dream-metaphysics.
		357	Only such a change will seriously modify the content of experience into which creation of objects made for use enters. And this modification of the nature of experience is the finally determining element in the esthetic quality of the experience of things produced.
Gallagher, Shaun; Zahavi, Dan	The Phenomenological Mind (2nd ed.) [98]	134	The point of departure has been the observation that it can often be quite difficult to distinguish a description of certain objects from a description of the experience of these very same objects. Back in 1903, G. E. Moore called attention to this fact, and dubbed it the peculiar diaphanous quality of experience: when you try to focus your attention on the intrinsic features of experience, you always seem to end up attending to what the experience is of.
Glynn, Ian	An Anatomy of Thought: The Origin and Machinery of Mind [104]	452	There are all sorts of other things we may be told about these experiences—what in the outside world is causing us to have them; what can be deduced about us or the outside world from the fact that we are having them; what effect they are having or likely to have on our behaviour; what is going on in our brains while we are having them –but none of this things tells us about the subjective qualities of the experiences; what it is like for us to have them.

Author(s)	Title [ref]	Page	Excerpt
Haybron, Daniel M.	The Pursuit of Unhappiness [116]	22	Indeed, it is surprisingly easy to err even about the present quality of one's experience, for instance because of the importance of elusive affects like anxiety for well-being.
		34	Crudely, hedonism identifies well-being with pleasure. A bit more precisely, well-being consists in a subject's balance of pleasant and unpleasant experience. The central idea is that what ultimately matters for welfare is the hedonic quality of individuals' experience, and nothing more.
		62	As to the nature of pleasure and unpleasure, hedonism comes in two basic flavors, which L. W. Sumner has called internalist and externalist. The former identifies pleasures in terms of a quality intrinsic to pleasurable experiences. On this sort of view, pleasure is kind of sensation, feeling, or quality of experience. For instance, what makes a given experience pleasant or unpleasant is a simple, unanalyzable feeling tone that it shares with all other such experiences.
		68	Similarly, an individual who is in depressed mood will likely find little pleasure in what happens, will tend to look on the dark side of things, and may more likely be saddened by negative events. Again, it is not clear what it could mean for someone to be in depressed mood if she lacked such propensities. This sort of disposition is not just a type or quality of experience.
		209	To the extent that we find it difficult to render judgments about how happy we are, or about the quality of our experience, we might rely somewhat on our expectations concerning how we should feel, or are likely to feel, given the circumstances.
Jay, Martin	Songs of Experience: Modern American and European Variations on a Universal Theme [154]	33	Among the consequences of the fetish method for the fortunes of "experience," four can be singled out for special attention. First, the new identification of reliable and certain experience with verifiable experimentation meant a belief in the repeatability and public quality of experience, at least when it was invoked to provide a source of valid knowledge.
		43	Whether it was reason or something else that flooded through the veins can be disputed, but Dilthey expressed a widespread assumption about the wan and restricted quality of experience examined by the figures he mentioned.
		103	In the latter, Rockefeller observes that "Dewey is careful to point out that his notion of the religious quality of experience does not refer to a special kind of experience that marked off from aesthetic, scientific, moral, or political experience or from experience as companionship and friendship. The religious quality of experience is not the result of interaction with some distinct religious object like a supernatural deity or the numinous" (p. 138).
		196	For a discussion of William's contribution to the idea of counter-history, which animated the New Historicist movement of literary criticism, see Catherine Gallagher and Stephen Greenblatt, *Practicing New Historicism* (Chicago, 2000), pp. 60-66. Gallagher and Greenblatt distinguish his work from Thompson's because of his greater sensitivity to the occluded, hidden, and repressed quality of experience.
Kahneman, Daniel	Thinking, Fast and Slow [160]	389	Finally, they indicated whether or not they intended to repeat or not to repeat the vacation they had just had. Statistical analysis established that the intentions for future vacations were entirely determined by the final evaluation—even when that score did not accurately represent the quality of the experience that was described in the diaries.

Author(s)	Title [ref]	Page	Excerpt
Kraut, Richard	What is Good and Why [179]	73	I think Nagel would have done better to drop the idea that the badness of the pain is immediately apprehended to say instead that the reason pain should be avoided (when it should be avoided) lies in something that every sufferer does apprehend: the quality of the experience, barely possible to describe, that we have when we are in pain.
		151	It would be a mistake, however, to suppose that their badness is to be explained solely by our reacting to them with dislike, and not even partially by the quality of the sensation to which we react with aversion. The quality of that experience—what nausea, for example, feels like—is what we react to when we try to avoid feeling it, and we justify our negative reaction on the basis of that sensation.
		152	Similarly, normal human beings react aversively to certain sensory stimuli, and their aversion is caused by the quality of the experience they have. There is no reason to set aside those experiences when we consider how well or badly their lives are going.
Levine, Robert	A Geography of Time [186]	47	Csikszentmihalyi has found, "hours seem to pass by in minutes, and occasionally a few seconds stretch out into what seems to be infinity. The clock no longer serves as a good analog of the temporal quality of experience.
McCarthy, John & Wright, Peter	Technology as Experience [201]	12	The lesson of the mobile phone and particularly of text messaging that seems not to have been learned yet is that the quality of experience is as much about the imagination of the consumers as it is about the product they are using.
		81	Compartmentalization of the senses is inconsistent with the quality of experience that Dewey argues is important for human growth and development.
Pilling, David	The Growth Delusion: The wealth and well-being of nations [235]	301	The story of the Eurostar with free Chateau Petrus hints at how wealthy societies may 'grow' in future by improving the quality of the experience. Quality—whether well prepared, locally grown food, personalized medical care, more cultural and outdoor activities, individually tailored products or better design—is lower carbon than quantity.
Porter, Michael	On Competition [242]	233	In tourism, for example, the quality of the visitor's experience depends not only on the appeal of the primary attraction (for example, beaches or historical sites) but also on the comfort and service of area hotels, restaurants, souvenir outlets, airport and other transportation facilities, and so on.
Prinz, Jesse	The Conscious Brain: How Attention Engenders Experience [243]	71	When we see a stick in water, it *looks* straight in some sense, but that doesn't mean that we have a straight visual image. Rather, the curve lines we experience happen to be what straight sticks look like when submerged, and we know this fact. Thus, "looks straight" expresses the representational content of our experience, not the qualities of the experience itself, but this way of talking makes it easy to confuse the content of experience for the quality.
		141	On the AIR theory, the spiking patterns give us the quality of experience, and the oscillations make those qualities conscious. [The AIR theory of Consciousness: Consciousness arises when and only when intermediate-level representations are modulated by attention.]

Author(s)	Title [ref]	Page	Excerpt
		151	Some authors have also claimed that we can experience objects (Clark, 1993), and others have suggested that we experience action affordances (Gibson, 1979; Noë, 2005). The crucial thing about these views is that they deny, implicitly or explicitly, that the quality of experience goes beyond these aspects of appearance.
		218	There is a thorny exegetical question about whether Kant's transcendental self is supposed to present itself somehow in the quality of experience. Since we do not know this self by description, it might be that there can be no feature of experience, no quale, that corresponds to it. On the other hand, Kant does imply that the quality of experience overall depends very much on this self (compare Kriegel, whose me-ness inheres in each feature of experience, because all sensory qualities are self-referential).
		290	But there is no evidence that these chemical differences make a difference for the quality of experience.
Siegel, Daniel	The mindful brain: Reflection and attunement in the cultivation of well-being [278]	69	When the richness of the textures of that bottom-up world become a part of our lives, it soon creates a quality of experience that lends itself to knowing when the seventh and eight sense are also felt with direct bottom-up simplicity. Sensing the mind's thoughts, feelings, memories, beliefs, attitudes, intentions, and perceptions is clearly less grounded in the physical world than our first six senses.
Sumner, L. W.	Welfare, Happiness & Ethics [292]	112	More importantly, however, hedonism underlines a truth which applies to all goods and ills, whether they consist merely our feelings or include states of the world. This truth is that nothing can make our lives go better or worse unless it somehow affects the quality of our experience.
Tversky, Amos	The Essential Tversky [308]	204	Is loss aversion irrational? This question raises a number of difficult normative issues. Questioning the values that decision makers assign to outcomes requires a criterion for the evaluation of preferences. The actual experience of consequences provides such a criterion: the value assigned to a consequence in a decision context can be justified as a prediction of the quality of the experience of that consequence (Kahneman & Snell, 1990).
Xiao, XiPeng	Technical, Commercial and Regulatory Challenges of QoS: An Internet Service Model Perspective [334]	14	People defining QoS this way acknowledge that reliability, security, routing policy, traffic engineering, etc. can all affect the end users' perception on network service quality. But they choose to take a "divide and conquer" approach in which "QoS" people will focus on traffic management, and other people will take care of reliability, security, routing, traffic engineering, etc. Together, QoS and other mechanisms will work to provide the Quality of Experience (QoE) to the end users. … End users care about network service quality, but they rarely distinguish between the terms QoS and QoE.

As to the books listed in Table A1.1., I would recommend the following five:

1. Dewey, J. (1934/2005). *Art as Experience.* Penguin, 363 p.
2. McCarthy, J. & Wright, P. (2004). *Technology as Experience*. MIT Press, 198 p.
3. Jay, M. (2006). *Songs of Experience: Modern American and European Variations on a Universal Theme*. University of California Press, 409 p.
4. Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan, 418 p.
5. Prinz, J. (2012). *The Conscious Brain*. Oxford University Press, 343 p.

Recommended articles

Anyone who wants to get familiar with the concept of quality of experience must read articles using the concept of quality of experience. To this end, this third part of the appendix presents a collection of quality of experience articles that I would like to recommend. The selected collection consists of 60 articles from different disciplines grouped into six sets of ten articles listed in Table A1.2. The first ten articles are, in my opinion, a good starting point, the next ten articles could be read next, and so on.[77]

Table A1.2: Recommended reading, 6 groups of 10 articles

Group 1

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE in the article
Barman, N., & Martini, M. G. (2019). QoE modeling for HTTP adaptive video streaming–a survey and open challenges. [30]	Networking, HAS	The main contribution of this paper is to present a comprehensive overview of recent and currently undergoing works in the field of QoE modeling for HAS.
Brooks, P., & Hestnes, B. (2010). User measures of quality of experience: why being objective and quantitative is important. [45]	Networking, measuring QoE	A structured approach to defining and measuring QoE in relation to QoS is explained and illustrated.
Brunnström, K., Beker, S. A., De Moor, K., Dooms, A., Egger, S., Garcia, M. N., ... & Lawlor, B. (2013). Qualinet white paper on definitions of quality of experience. [47]	Networking, QoE definitions	The Network of Excellence (NoE) Qualinet aims at extending the notion of network-centric Quality of Service (QoS) in multimedia systems, by relying on the concept of Quality of Experience (QoE).
Chikkerur, S., Sundaram, V., Reisslein, M., & Karam, L. J. (2011). Objective video quality assessment methods: A classification, review, and performance comparison. [57]	HCI, video quality	Controlling and monitoring the QoS parameters of the individual system components by appropriately selecting system parameters (such as compression ratios and reserved network bandwidth) is important for efficiently achieving high overall system performance and user QoE.
Csikszentmihalyi, M., & LeFevre, J. (1989). Optimal experience in work and leisure. [66]	Psychology, work and leisure	Regardless of the quality of experience, however, respondents are more motivated in leisure than in work.
Fiedler, M., Hoßfeld, T., & Tran-Gia, P. (2010). A generic quantitative relationship between quality of experience and quality of service. [90]	Networking, QoS vs. QoE	Against this background, this article proposes a generic formula in which QoE and QoS parameters are connected through an exponential relationship, called IQX hypothesis.
Kahneman, D. (1999). Objective happiness. [161]	Psychology, well-being	One suggestion has already been mentioned on several occasions: because immediate reports of the quality of experience avoid the difficulties of memory and of integration, experience-sampling methods have significant advantages and should be used whenever possible (Stone et al., this volume).
Otto, J. E., & Ritchie, J. B. (1996). The service experience in tourism. [228]	Tourism, service experience	Comparison of QOS and QOE frameworks.

[77] For this purpose, I read 104 quality of experience articles. Based on the reading experience, the average reading time is 1.5 hours per article, although there is a lot of variation in the length and degree of difficulty of the articles.

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE in the article
Shin, D. (2019). How does immersion work in augmented reality games? A user-centric view of immersion and engagement. [276]	HCI, augmented reality	When users verify the quality and obtain satisfaction from the features, it is referred to as quality of experience.
Wechsung, I. & De Moor, K. (2014). Quality of experience versus user experience. [326]	HCI, QoE vs. user experience	Thus the literature within the User Experience domain can be of great value for the Quality of Experience-community, especially if the latter intends to really put the recently proposed more holistic definition of Quality of Experience into practice.

Group 2.

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE
Balachandran, A., Sekar, V., Akella, A., Seshan, S., Stoica, I., & Zhang, H. (2013). Developing a predictive model of quality of experience for internet video. [22]	Networking, Internet video	The goal of this paper is to develop a predictive model of Internet video QoE.
Baloglu, S., & McCleary, K. W. (1999). A model of destination image formation. [24]	Tourism, destination image	Moderate support was provided for this hypothesis as quality of experience (COG1=.073) and value/environment (COG3=.133) positively influenced overall image (OI).
Ellis, G. D., Freeman, P. A., Jamal, T., & Jiang, J. (2019). A theory of structured experience. [83]	Tourism, structured experience	In evaluated experience models, the quality of experience may be generalized across diverse point-of-service encounters for an entire visit to a destination or the duration of participation in an event (e.g. Aho 2001; Kao, Huang, and Wu 2008; Mossberg 2007; Otto and Ritchie 1996).[78]
Engelke, U., Darcy, D. P., Mulliken, G. H., Bosse, S., Martini, M. G., Arndt, S., ... & Brunnström, K. (2017). Psychophysiology-based QoE assessment: A survey. [86]	HCI, psycho-physiology	We summarize multimodal techniques and discuss several important aspects of psychophysiology-based QoE assessment, including the synergies with psychophysical assessment and the need for standardized experimental design.
Forlizzi, J., & Ford, S. (2000). The building blocks of experience: an early framework for interaction designers. [95]	HCI, interaction design	In 1994, a panel of judges for interactions magazine took a step towards clarifying user-product experience, by creating a set of criteria for assessing qualities of experience of entries for an ACM-sponsored design contest.
Hoßfeld, T., Egger, S., Schatz, R., Fiedler, M., Masuch, K., & Lorentzen, C. (2012). Initial delay vs. interruptions: Between the devil and the deep blue sea. [125]	Networking, delay vs. interruption	Insufficient resources (e.g. low transmission capacity), network problems (e.g. high latency), or time-consuming operations (e.g. user authentication, Internet connection setup) open a plethora of design options how to deal with these QoE impairments – typically translated into waiting times for the end user.

[78] This is the only sentence in which quality of experience is used in the paper. However, the paper provides valuable insight into experiences in general.

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE
Reichl, P., Tuffin, B., & Schatz, R. (2013). Logarithmic laws in service quality perception: where microeconomics meets psychophysics and quality of experience. [250]	Networking, perception of quality	While the discussion about how to define QoE properly is still ongoing, the focus of current research is increasingly directed towards defining reliable and reproducible quantitative metrics which link together technical system parameters with the perceptual quality of the user.
Schiefele, U. (1991). Interest, learning, and motivation. [263]	Psychology, student motivation	The quality of experience in the classroom was measured by means of the Experience Sampling Method (ESM; Csikszentmihalyi & Larson, 1987).
Seufert, M., Egger, S., Slanina, M., Zinner, T., Hoßfeld, T., & Tran-Gia, P. (2015). A survey on quality of experience of HTTP adaptive streaming. [266]	Networking, HAS	The technical development of HAS, existing open standardized solutions, but also proprietary solutions are reviewed in this paper as fundamental to derive the QoE influence factors that emerge as a result of adaptation.
Suri, J. F. (2003). The experience of evolution: developments in design practice. [293]	HCI, design practices	Today's designers and clients are concerned with the quality of experiences people will have.

Group 3.

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE
Baraković Husić, J., Baraković, S., Cero, E., Slamnik, N., Oćuz, M., Dedović, A., & Zupčić, O. (2020). Quality of experience for unified communications: A survey. [29]	Networking, influencing factors	The results of the qualitative review include various IFs, QoE dimensions, and key findings in the form of research recommendations for QoE in the context of UC.
Csikszentmihalyi, M. (2000). The costs and benefits of consuming. [65]	Psychology, consumption	Yet trying to fill unstructured time with passive entertainment does not work well; the quality of experience while watching TV is barely more positive than that of the slough of despond that awaits the unfocused mind (Kubey and Csikszentmihalyi 1990).
Domínguez-Quintero, A. M., González-Rodríguez, M. R., & Paddison, B. (2020). The mediating role of experience quality on authenticity and satisfaction in the context of cultural-heritage tourism. [79]	Tourism, authenticity & satisfaction	The findings also identified the mediating role of quality of experience on authenticity and satisfaction.
Harman, G. (1990). The intrinsic quality of experience. [112]	Philosophy, conscious experience	First, when you attend to a pain in your leg or to your experience of the redness of an apple, you are aware of an intrinsic quality of your experience, where an intrinsic quality is a quality something has in itself, apart from its relations to other things.
Lallemand, C., Gronier, G., & Koenig, V. (2015). User experience: A concept without consensus? Exploring practitioners' perspectives through an international survey. [182]	HCI, user experience	Similarly, respondents deplore the fact that D5 defines UX by the too generic wording 'quality of experience', which is actually a statement of the obvious.

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE
Raake, A., & Egger, S. (2014). Quality and quality of experience. [244]	Networking, quality	The present chapter lays out the basis for the understanding of Quality of Experience (QoE) as it is followed by the book.
Reiter, U., Brunnström, K., De Moor, K., Larabi, M. C., Pereira, M., Pinheiro, A., ... & Zgank, A. (2014). Factors influencing quality of experience. [251]	HCI, influencing factors	With respect to Human IFs, we discuss variant and stable factors that may potentially bear an influence on QoE, either for low-level (bottom-up) or higher-level (top-down) cognitive processing.
Streijl, R. C., Winkler, S., & Hands, D. S. (2016). Mean opinion score (MOS) revisited: methods and applications, limitations and alternatives. [290]	Networking, MOS	From a service provider's perspective, "maximizing QoE" may have different objectives; these could be maximizing overall QoE for multiple users in a network, maximizing the QoE of certain individual users or groups, maximizing the number of "satisfied" users, etc.
Takahashi, A., Hands, D., & Barriac, V. (2008). Standardization activities in the ITU for a QoE assessment of IPTV. [294]	Networking, standardization	Industry must have access to tools designed to assess the QoE of IPTV services.
Van Moorsel, A. (2001). Metrics for the Internet Age: Quality of Experience and Quality of Business. [312]	Networking, business	It builds up an evaluation framework for Internet services, relating QoS (quality of service), QoE (quality of experience) and QoBiz.

Group 4.

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE
Akhtar, Z., & Falk, T. H. (2017). Audio-visual multimedia quality assessment: A comprehensive survey. [7]	HCI, audio-visual quality	Multimedia service providers are formulating various techniques to provide better quality of experience (QoE), which is increasingly being demanded by end-users.
Antonakoglou, K., Xu, X., Steinbach, E., Mahmoodi, T., & Dohler, M. (2018). Toward haptic communications over the 5G tactile Internet. [14]	Networking, haptic	Since evaluating QoE in haptic-based applications with force feedback over the Internet is a process that has only recently taken its first steps, the way to resolve this open issue is still under investigation.
Crites, S. (1971). The narrative quality of experience. [61]	Philosophy, experiences as stories	So the narrative quality of experience has three dimensions, the sacred story, the mundane stories, and the temporal form of experience itself: three narrative tracks, each constantly reflecting and affecting the course of the others.
Laghari, K. U. R., Crespi, N., and Connelly, K. (2012). Toward total quality of experience: A QoE model in a communication ecosystem. [181]	Networking, ecosystem	Quality of experience (QoE) is a fast emerging multidisciplinary field based on social psychology, cognitive science, economics, and Engineering science, focused on understanding overall human quality requirements.
Mok, R. K., Chan, E. W., & Chang, R. K. (2011). Measuring the quality of experience of HTTP video streaming. [212]	Networking, measuring QoE	Our ultimate goal is to understand how the network QoS affects the QoE of HTTP video streaming.

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE
Pilloni, V., Floris, A., Meloni, A., & Atzori, L. (2018). Smart home energy management including renewable sources: A qoe-driven approach. [236]	Energy, power management	The assigned profile is then exploited by the QoE-aware cost saving appliance scheduling and the QoE-aware renewable source power allocation algorithms.
Riva, G., Banos, R. M., Botella, C., Wiederhold, B. K., & Gaggioli, A. (2012). Positive technology: using interactive technologies to promote positive functioning. [252]	Psychology, positive technologies	In this paper, we suggest that it is possible to use technology to manipulate the quality of experience, with the goal of increasing wellness, and generating strengths and resilience in individuals, organizations, and society.
Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. [254]	Psychology, motivation	Over three decades of research has shown that the quality of experience and performance can be very different when one is behaving for intrinsic versus extrinsic reasons.[79]
Tian-Cole, S., Crompton, J. L., & Willson, V. L. (2002). An empirical investigation of the relationships between service quality, satisfaction and behavioral intentions among visitors to a wildlife refuge. [302]	Tourism, behavioral intention	At the transaction level, the concepts of quality of performance and quality of experience are conceptualized as direct antecedents of overall service quality and overall satisfaction.
Zhu, Y., Heynderickx, I., & Redi, J. A. (2015). Understanding the role of social context and user factors in video quality of experience. [351]	Networking, video quality	However, recent studies have shown that this approach cannot sufficiently estimate user satisfaction, and that QoE depends on multiple factors, besides the media technical properties.

Group 5.

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE
Alben, L. (1996). Defining the criteria for effective interaction design: Quality of experience. [9]	HCI, design	We call this "quality of experience."[80]
Casas, P., D'Alconzo, A., Wamser, F., Seufert, M., Gardlo, B., Schwind, A., ... & Schatz, R. (2017). Predicting QoE in cellular networks using machine learning and in-smartphone measurements. [50]	Networking, machine learning	We conceive different QoE assessment models based on supervised machine learning techniques, which are capable to predict the QoE experienced by the end user of popular smartphone apps (e.g., YouTube and Facebook), using as input the passive in-device measurements.
Chen, M., Ma, Y., Li, Y., Wu, D., Zhang, Y., & Youn, C. H. (2017). Wearable 2.0: Enabling human-cloud integration in next generation healthcare systems. [55]	HCI, healthcare	Therefore, in this article we propose a Wearable 2.0-based healthcare services to improve quality of experience (QoE) and quality of service (QoS) in the next generation healthcare system.
Chen, Y., Wu, K., & Zhang, Q. (2015). From QoS to QoE: A tutorial on video quality assessment. [56]	Networking, video quality	Traditionally, QoE is obtained from subjective test, where human viewers evaluate the quality of tested videos under a laboratory environment.

[79] This is the only sentence in which quality of experience appears in the paper.
[80] Surprisingly, this is the only sentence in which quality of experience is used in the paper. The paper uses various other closely related concepts, like user experience, quality experience, and quality user experience. Quality of experience appears, however, in the title of the paper as well as in a figure.

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE
Egger, S., Hoßfeld, T., Schatz, R., & Fiedler, M. (2012). Waiting times in quality of experience for web based services. [82]	Networking, waiting	It then investigates to which extent the same relationships can also be used to describe QoE for more complex services such as web browsing.
Jin, N., Lee, S., & Lee, H. (2015). The effect of experience quality on perceived value, satisfaction, image and behavioral intention of water park patrons: New versus repeat visitors. [156]	Tourism, behavioral intention	These complex interrelationships between quality of experience at a water park and direct (e.g. perceived value, image and satisfaction) and indirect outcomes (e.g. behavioral intention) create an imperative to improve insight into the effects of new attractions on the commercial performance of an enterprise involving a water park.
Melzack, R. (1993). Pain: past, present and future. [204]	Psychology, pain	For example, when we respond to the experience of pain or itch, it is evident that the experience has been synthesized by the body-self neuromatrix (or relevant neuromodules) sufficiently for the neuromatrix to have imparted the neurosignature patterns that underlie the quality of experience, affect and meaning.
Schatz, R., Egger, S., & Platzer, A. (2011). Poor, good enough or even better? Bridging the gap between acceptability and QoE of mobile broadband data services. [261]	Networking, acceptability	To this end, we perform a critical review of the term's utilization and operationalization in related fields and discuss the results of a series of mobile broadband QoE studies performed in lab and field settings.
Tononi, G., & Koch, C. (2015). Consciousness: here, there and everywhere? [304]	Philosophy, consciousness	The 'form' or shape of the quale (constellation of stars) is identical to the quality of the experience.
Wu, W., Arefin, A., Rivas, R., Nahrstedt, K., Sheppard, R., & Yang, Z. (2009). Quality of experience in distributed interactive multimedia environments: toward a theoretical framework. [333]	HCI, interactive multimedia	The results present the first deep study to model the multi-facet QoE construct, map the QoS-QoE relationship, and capture the human-centric quality modalities in the context of DIMEs.

Group 6.

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE
Hoßfeld, T., Keimel, C., Hirth, M., Gardlo, B., Habigt, J., Diepold, K., & Tran-Gia, P. (2013). Best practices for QoE crowd-testing: QoE assessment with crowdsourcing. [127]	Networking, crowd-testing	The advantages of QoE crowd-testing lie not only in the reduced time and costs for the tests, but also in a large and diverse panel of international, geographically distributed users in realistic user settings.
Kahneman, D., & Snell, J. (1992). Predicting a changing taste: Do people know what they will like? [163]	Psychology, experienced and decision utility	Little is left in this usage of the original sense of utility in the writings of Bernoulli, Bentham, and Mills, where it was identified with the hedonic quality of experience (Stigler, 1950).[81]

[81] This is the only place in which quality of experience appears in the paper. However, this paper describes the concepts of experienced utility and decision utility vital for modeling the value and usefulness of experiences.

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE
Kivela, J., & Crotts, J. C. (2006). Tourism and gastronomy: Gastronomy's influence on how tourists experience a destination. [177]	Tourism, gastronomy	Regression analysis, dependent variable: Overall, what contribution did food and cuisine have on your overall quality of experience as a tourist visiting Hong Kong?
Mao, H., Netravali, R., & Alizadeh, M. (2017). Neural adaptive video streaming with pensieve. [196]	Networking, HAS	As a result, Pensieve automatically learns ABR algorithms that adapt to a wide range of environments and QoE metrics.
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. [219]	HCI, acceptance	Findings are also interpreted within the context of IT and cognitive/behavioral science perspectives, further providing for face validity of the quality of experience construct.
Möller, S., Engelbrecht, K. P., Kuhnel, C., Wechsung, I., & Weiss, B. (2009). A taxonomy of quality of service and quality of experience of multimodal human-machine interaction. [214]	HCI, multimodal interaction	It consists of three layers: (1) The QoS-influencing factors related to the user, the system, and the context of use; (2) the QoS interaction performance aspects describing user and system behavior and performance; and (3) the QoE aspects related to the quality perception and judgment processes taking place inside the user.
Perkis, A., Timmerer, C., Baraković, S., Husić, J. B., Bech, S., Bosse, S., ... & Zadtootaghaj, S. (2020). QUALINET white paper on definitions of immersive media experience (IMEx). [233]	HCI, immersive media	In this section, we will establish a relationship between the concepts of Quality of Experience (QoE) and immersive media experience (IMEx).
Shin, D. (2018). Empathy and embodied experience in virtual environment: To what extent can virtual reality stimulate empathy and embodied experience? [274]	HCI, virtual environment	It has also explored the two-tiered process of immersion that includes user experience (presence and flow) and quality of experience (empathy and embodiment).

Authors (year), title, [ref]	Discipline, topic	Third sentence with quality of experience or QoE
Winkler, S., & Mohandas, P. (2008). The evolution of video quality measurement: From PSNR to hybrid metrics. [330]	Networking, measurement	Unfortunately, quality in this context is a rather ill-defined concept—we list just some of the numerous factors contributing to QoE here [2]–[4]:[82] • Individual interests of the viewer, such as favorite programs, which determine the level and focus of attention; • Quality expectations of the viewer, for example a feature film screened in a cinema vs. a short clip watched on a mobile device; • Video experience of the viewer, which also determines quality expectations (once you have seen high-definition content it's hard to go back); • Display type (CRT, LCD, etc.) and properties (size, resolution, brightness, contrast, color, response time); • Viewing setup and conditions, such as viewing distance or ambient/exterior light; • Quality and synchronization of the accompanying audio; • Interaction with the service or display device (e.g. zap time, remote control, electronic program guide).
Zhao, T., Liu, Q., & Chen, C. W. (2017). QoE in video transmission: A user experience-driven strategy. [348]	Networking, video quality	This advantage has raised the popularity and widespread usage of QoE in video transmission.

[82] The numbers [2]-[4] refer to references in Winkler & Mohandas [330]. None of those three references uses the term quality of experience but either image quality or experienced quality.

Appendix 2. The number of the quality of experience papers

The apparent objective of this appendix is to estimate the number of the quality of experience papers in different fields of study. Moreover, the distribution tells something about the nature of different research fields; in some fields, a few authors and papers dominate the literature while some other fields are more decentralized. These differences are analyzed by a long tail model. The formula used to estimate the complete citation distributions for different fields of study is the following [92, 172]:

$$F(k, \alpha, \beta, N_{50}) = \frac{\beta}{1 + \left(\frac{N_{50}}{k}\right)^{\alpha}}$$

where:

$F(k)$ = the popularity covered by items up to rank *k*,
N_{50} = the number of items that cover half of the total popularity or volume,
α = factor defining the form of the function, and
β = total volume of all items.

The popularity of k[th] item is then calculated as the difference between two consecutive cumulative values: $f(k) = F(k) - F(k-1)$. The share of the most popular item $f(1)$ is defined as $F(1)$.[83]

As regards the citation distributions, networking and HCI fields form a distinct group compared to other fields. Figure A2.1 shows three long tail distributions: networking papers, HCI papers, and all other papers. The ranking implies that the first item (rank = 1) is the paper with the largest number of citations, the second item (rank = 2) is the paper with the second largest number of citations, etc. For instance, Agiwal et al. [5] with 2255 citations is the ranking number one in the network area, Fiedler et al. [90] with 983 citations is the second, and Jiang et al. [155] with 908 citations is the third (these are the first three dots in the series "network, data"). Thus, the horizontal axis indicates the rank number, and the vertical axis indicates the number of citations of the paper with the given rank. Note that both axes are shown on a logarithmic scale.

The main feature visible in Figure A2.1 is the difference between technical and non-technical fields. First, the number of networking papers in Set Q3 is much larger than the number of all non-technical papers combined. Second, the citations are much more concentrated among non-technical papers than among technical papers. The most cited networking paper in Set Q3, [5] with 2255 citations, represents about two percent of the estimated total number of quotes on networking papers (126 715 on the row β in Table A2.1.). In contrast, the most cited non-technical ('others') paper in Set Q3, [24] with 5100 citations, represents about eight percent of the total number of quotes on non-technical papers in Set Q3. Moreover, the long-tail model results in an estimate that only 18 non-technical papers are needed to cover half of the quotes on non-technical papers in Set Q3. The corresponding number is almost 600 for networking papers (shown on the row N_{50} in Table A2.1.). Therefore, the quality of experience research in the networking field is more fragmented than the research on other fields of research.

In Set Q3, the long tail model intersects with the level of one citation roughly at 2000 for HCI papers, 10000 for networking papers, and 1400 for other papers. These numbers provide estimates for the total number of papers meeting the criteria of one citation in each of the sets. However, several weaknesses in the data sets reduce the accuracy of the estimate. Above all, the coverage of papers (meeting the criterion of one citation) is better in HCI and networking than in philosophy, psychology, and tourism.[84] Moreover, there is no guarantee that the true distribution conforms the long tail model utilized in the analysis.

Set Q3C covers approximately 85% of papers in HCI and networking fields whereas in other fields the coverage is lower, perhaps 70%. Thus, the total number of papers with at least three appearances of quality of experience

[83] A consequence of this definition, $f(1) = F(1)$, is that the first item is much more popular than the second item when α is small, for instance, in the case of HCI papers (where $\alpha = 0.545$). Although this property may seem arbitrary, many examples of real distributions support this assumption (see also examples in [92] and [172]).

[84] The difference can be partly explained by the fact that Aalto University has broad agreements with publishers in the field of networking and HCI, but less comprehensive agreements with publishers in many other fields of research.

and with at least one citation was roughly 16000 in September 2021.[85] This estimation can be compared to the number given Google Scholar as to the total number of publications with "quality of experience," 31500.[86] Because many papers are not citated at all and many papers contain only one or two occasions of quality of experience, these two estimates, 16000 and 31500, are not inconsistent with each other.

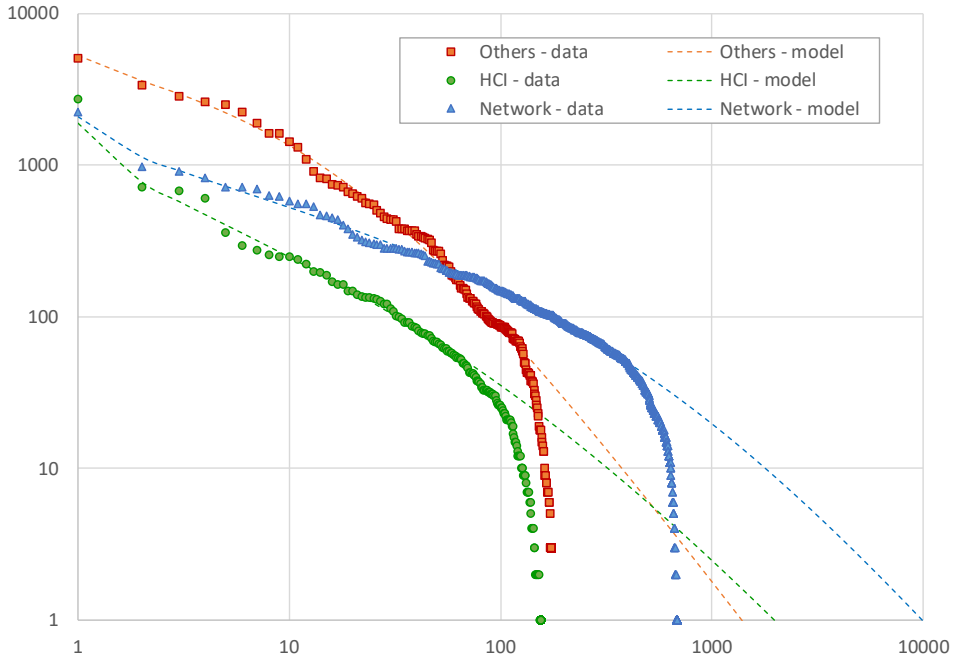

Figure A2.1: Long tails of citations for papers with at least three mentions of quality of experience for networking papers, HCI papers, and other papers (in Set Q3). Dotted (slightly curved) lines present a long tail model fitted to the citation data above 70 citations. The first paper is not included in the fitting.

[85] 1400/0.7 + 2000/0.85 + 10000/0.85 = 16118

[86] Data were collected on June 8, 2022. Note also that the number given by Google Scholar should not be considered as an exact number of citations but as a more or less accurate estimate.

Table A2.1: The number of citations to papers in Set Q3 in rank order (black) and respective long tail models (blue). "Others" include all papers that are not classified into HCI or network fields. k:(f=1) provides an estimate of the total number of papers with at least one citation.

	Others Data	Others Model	HCI Data	HCI Model	Network Data	Network Model
N_{50}		18.2		105.4		595.9
α		0.839		0.545		0.639
β		65500		25752		126715
k:(f=1)		1379		1968		9890
1	5100	5280.4	2740	1888.8	2255	2096.6
2	3360	3602.9	714	776.5	983	1139.1
3	2838	2951.1	679	575.7	908	926.0
4	2624	2523.9	603	470.5	820	807.3
5	2520	2209.7	359	402.9	717	727.6
6	2259	1964.8	296	354.9	713	668.8
7	1898	1766.8	274	318.6	702	622.9
8	1632	1602.7	257	289.9	635	585.5
9	1613	1464.2	251	266.6	627	554.3
10	1423	1345.6	251	247.1	582	527.6
20	647	704.4	149	146.6	348	377.2
30	439	447.2	114	105.6	284	305.9
40	352	313.5	84	82.7	262	261.4
50	271	233.8	68	67.9	224	230.2
60	185	182.0	56	57.5	194	206.6
70	134	146.3	46	49.8	186	187.9
80	109	120.4	36	43.8	174	172.7
90	93	101.1	31	39.1	163	160.0
100	86	86.2	25	35.2	148	149.1
120	69	65.1	13	29.3	132	131.6
150	23	45.8	2	23.2	111	112.1
200		28.7		17.1	90	90.0
300		14.6		10.8	66	64.5
400		8.9		7.7	49	49.9
500		6.0		5.9	31	40.4
600		4.4		4.7	16	33.8
1000		1.8		2.5		19.7
1400		**1.0**		1.6		13.4
2000		0.5		**1.0**		8.8
10000		0.0		0.1		**1.0**

Appendix 3. Vocabulary in quality of experience papers

In this appendix, the content of the papers in different fields of research is compared using the following process. First, I took all 1026 papers in Set Q3 and picked the first three sentences in which either quality of experience or QoE appeared. Then, I counted the number of different words in seven sets of papers: 1) philosophical papers, 2) psychological papers, 3) tourism papers, 4) HCI papers, 5) networking papers published between 2000 and 2010, 6) networking papers published between 2011 and 2016, and 7) networking papers published between 2017 and 2020.

Then, I identified the six most special words in each set by calculating a simple index for the m:th word in the n:th category as follows:

$$I(m,n) = \frac{N(m,n)}{\sum_{i=1}^{199} N(i,n)} \cdot \frac{\sum_{j=1}^{7}\sum_{i=1}^{199} N(i,j)}{\sum_{j=1}^{7} N(m,j)}$$

where $N(m,n)$ is the number of instances of the m:th word in the n:th category. Finally, I ordered the words in each category (n) according to the index $I(m,n)$. The ten most characteristic words in each category are presented in Table A3.1. The whole data for the 199 most frequent words are shown in Table A3.3. Note that singular and plural forms and different tenses are counted separately.

Table A3.1: The ten most characteristic words in seven sets of papers.

	Philosophy	Psychology	Tourism	HCI	Network (-2010)	Network (2011-16)	Network (2017-20)
1.	value	students	value	human	IPTV	adaptation	computing
2.	what	were	satisfaction	factors	parameters	access	latency
3.	through	was	experiences	aspects	loss	wireless	improve
4.	conditions	learning	environment	study	control	optimization	problem
5.	its	experiences	were	understanding	measurement	cloud	videos
6.	there	they	was	context	packet	resource	adaptive
7.	it	process	overall	influence	QoS	aware	delay
8.	or	challenges	significant	media	framework	optimal	communication
9.	but	each	influence	related	use	traffic	real
10.	information	study	at	design	service	architecture	cloud

The most surprising observation is that the words *was* and *were* are much more common in non-technical fields than in the technical fields of HCI and networking. It seems that in non-technical disciplines, researchers conduct studies and then report their conclusions. Researchers have problems in technical areas and then describe what their solutions are for the problems. "Engineers like problems they can solve," as Walter Vincenti puts it according to W. Brian Arthur [18, p. 15]. In contrast, scientists like peculiar phenomena for which they can give credible explanations.

Table A3.1. also demonstrates the changes in the research topics in the networking field. In the early phase before 2010, the main topic was the control and effect of packet losses in IPTV systems. In the middle phase, the main task was resource optimization in the wireless access networks. In the recent years, the main objective has been to improve the algorithms for adaptive video streaming.

Table A3.2. provides additional insight into the differences between different disciplines through the correlation between the categories shown in Table A3.3. However, the correlation is not calculated between the actual numbers, but between logarithmic values, as follows:

$$M(m,n) = \log_{10}\big(1 + N(m,n)\big)$$

The reason to use logarithmic values is to avoid a situation in which the most common words dominate the calculation of correlations.

The main observations are the following. The vocabularies in philosophy, psychology, and tourism are similar but not identical. In contrast, the vocabulary used in non-technical fields differ essentially from the vocabulary used in the area of networking. HCI locates between networking and non-technical fields. Based on the correlations, HCI is somewhat closer to networking than to the non-technical fields. The vocabulary in the early phase of networking research resembled a little more the vocabulary in HCI and non-technical fields than what the vocabulary used in more recent years resembles.

Table A3.2: Correlation in vocabulary between different fields and different periods in the case of networking research. Correlations are calculated by the logarithmic values, $M(m,n)$.

	Philosophy	Psychology	Tourism	HCI	Network (-2010)	Network (2011-16)	Network (2017-20)
Philosophy		0,722	0,721	0,656	0,556	0,538	0,518
Psychology	0,722		0,752	0,663	0,491	0,405	0,404
Tourism	0,721	0,752		0,687	0,592	0,525	0,500
HCI	0,656	0,663	0,687		0,794	0,771	0,755
Network (-2010)	0,556	0,491	0,592	0,794		0,849	0,801
Network (2011-16)	0,538	0,405	0,525	0,771	0,849		0,940
Network (2017-20)	0,518	0,404	0,500	0,755	0,801	0,940	

Table A3.3: List of words with at least 50 appearances in the first three quality of experience sentences. Note that singular and plural forms (e.g., experience and experiences) are counted separately. The six most characteristic words (per column) are marked by red color.

Word	Philosophy	Psychology	Tourism	HCI	Network (-2010)	Network (2011-16)	Network (2017-20)	**In total**
In total	2138	7903	3331	12443	12443	22515	22642	**83415**
the	187	568	233	825	853	1393	1339	**5398**
of	161	564	241	808	743	1230	1199	**4946**
and	59	293	147	414	352	705	709	**2679**
QoE	0	14	11	371	458	773	784	**2411**
quality	65	269	138	386	376	558	525	**2317**
to	54	169	73	294	402	592	629	**2213**
experience	77	313	116	317	257	408	407	**1895**
in	38	201	56	255	232	437	434	**1653**
a	54	140	74	243	276	460	395	**1642**
is	50	99	37	159	188	280	288	**1101**
for	11	65	20	141	169	321	318	**1045**
user	0	2	2	130	144	235	230	**743**
on	13	60	37	128	121	186	177	**722**
as	10	71	46	105	117	174	167	**690**
that	41	95	29	104	95	156	162	**682**
this	16	39	22	110	128	189	161	**665**
video	0	2	0	57	82	237	186	**564**
we	4	30	2	91	90	171	145	**533**
by	23	53	26	73	99	135	114	**523**
service	0	1	17	57	144	149	137	**505**
with	9	81	8	64	58	123	161	**504**
network	0	0	0	14	102	203	167	**486**
users	0	1	3	54	69	173	180	**480**

Word	Philosophy	Psychology	Tourism	HCI	Network (-2010)	Network (2011-16)	Network (2017-20)	In total
be	18	29	14	50	82	99	86	**378**
an	15	27	22	60	74	89	90	**377**
QoS	0	2	2	40	132	93	106	**375**
are	13	29	14	58	42	86	93	**335**
which	20	29	9	44	46	81	85	**314**
end	2	2	0	28	81	97	84	**294**
it	22	30	9	44	44	68	75	**292**
or	19	54	15	33	43	59	55	**278**
from	8	21	15	43	37	89	60	**273**
paper	0	7	7	41	57	92	63	**267**
services	0	0	2	31	71	70	88	**262**
can	9	10	8	32	40	67	87	**253**
based	1	9	3	33	38	70	94	**248**
has	5	12	6	39	39	78	63	**242**
such	4	3	6	42	26	62	69	**212**
mobile	0	0	0	14	19	89	85	**207**
multimedia	0	3	0	40	44	59	54	**200**
content	4	0	0	40	29	55	65	**193**
applications	0	3	0	27	33	57	67	**187**
streaming	0	0	0	15	24	65	81	**185**
different	1	11	8	30	28	61	44	**183**
perceived	0	20	11	25	42	50	35	**183**
performance	0	7	14	27	37	46	52	**183**
high	0	21	0	13	26	45	77	**182**
networks	0	1	0	5	32	73	70	**181**
at	0	14	20	19	27	45	52	**177**
subjective	0	11	3	35	34	55	32	**170**
have	5	22	8	29	19	46	39	**168**
their	4	17	8	25	16	48	48	**166**
time	3	18	8	14	24	42	54	**163**
these	4	17	5	35	24	39	37	**161**
not	7	20	7	29	25	34	38	**160**
system	2	3	0	33	27	35	53	**153**
between	6	27	14	19	24	36	25	**151**
model	0	5	8	31	31	45	30	**150**
application	0	3	0	16	42	57	28	**146**
our	1	10	3	25	22	42	42	**145**
how	1	10	5	29	27	33	37	**142**
factors	0	13	6	**51**	15	34	17	**136**
data	0	2	1	9	21	45	57	**135**
more	3	24	7	15	20	38	25	**132**
also	4	16	9	23	17	27	32	**128**
level	1	9	10	14	27	31	31	**123**
results	1	9	5	27	14	35	32	**123**
study	0	28	8	**40**	9	18	20	**123**
improve	0	0	2	8	9	35	**68**	**122**

Word	Philosophy	Psychology	Tourism	HCI	Network (-2010)	Network (2011-16)	Network (2017-20)	In total
parameters	0	0	4	9	51	33	23	**120**
provide	0	4	2	12	29	36	36	**119**
using	1	11	1	28	22	27	29	**119**
been	3	11	4	24	12	32	32	**118**
over	0	8	0	6	22	40	42	**118**
metrics	0	0	0	11	24	40	41	**116**
while	1	10	7	14	16	33	35	**116**
overall	0	12	14	24	19	26	17	**112**
proposed	0	2	2	16	16	28	48	**112**
research	0	8	5	24	20	33	22	**112**
into	2	8	2	19	23	32	23	**109**
however	1	6	0	17	17	32	35	**108**
requirements	0	2	1	6	20	32	43	**104**
when	4	24	0	23	12	19	22	**104**
approach	2	1	5	14	19	37	25	**103**
context	3	7	2	34	11	26	20	**103**
resource	0	0	0	1	11	49	42	**103**
satisfaction	0	2	36	7	10	22	26	**103**
wireless	0	0	0	1	12	53	36	**102**
but	7	7	6	18	12	32	18	**100**
important	0	11	1	17	20	24	27	**100**
objective	0	4	0	20	21	27	28	**100**
both	6	14	2	18	10	20	29	**99**
impact	0	2	2	18	17	29	30	**98**
order	1	5	2	14	21	27	27	**97**
terms	3	7	5	11	10	35	25	**96**
new	0	4	1	16	24	24	26	**95**
propose	0	4	0	12	11	38	30	**95**
traffic	0	0	0	5	8	43	39	**95**
its	8	6	4	14	17	19	26	**94**
management	0	0	2	5	24	29	34	**94**
systems	4	0	0	24	14	27	25	**94**
used	1	7	0	23	17	27	19	**94**
work	3	17	0	11	14	23	25	**93**
design	0	1	0	27	23	19	22	**92**
may	4	11	2	17	14	22	22	**92**
assessment	0	3	2	26	16	25	19	**91**
delivery	0	0	0	16	15	36	24	**91**
learning	0	43	0	12	3	7	26	**91**
will	6	4	2	18	14	19	28	**91**
well	1	13	3	23	14	24	12	**90**
evaluation	0	5	1	21	14	29	19	**89**
rate	0	2	1	4	14	39	29	**89**
Internet	0	2	0	4	18	36	27	**87**
perception	4	2	2	24	20	21	14	**87**

Word	Philosophy	Psychology	Tourism	HCI	Network (-2010)	Network (2011-16)	Network (2017-20)	In total
was	0	38	14	13	8	5	9	87
e.g.	0	0	3	13	9	36	25	86
human	1	5	3	35	11	14	17	86
other	5	14	3	15	6	20	22	85
they	2	25	5	17	11	12	13	85
related	1	12	1	23	16	20	11	84
bandwidth	0	0	0	9	7	31	35	82
use	1	10	5	15	21	18	12	82
cloud	0	0	0	5	0	39	37	81
influence	2	10	8	23	7	18	13	81
providers	0	0	0	12	13	29	27	81
one	5	4	2	19	13	21	16	80
framework	0	5	1	11	22	15	22	76
studies	0	11	3	15	4	23	20	76
control	0	8	0	5	26	18	18	75
problem	0	1	0	7	3	25	39	75
through	8	7	6	8	13	14	19	75
each	2	20	5	7	7	21	12	74
experiences	3	23	14	15	7	7	5	74
delay	0	0	0	8	8	23	34	73
latency	0	0	0	6	4	16	47	73
were	0	41	11	11	3	2	5	73
adaptation	0	0	0	9	11	39	13	72
measure	0	3	2	16	15	25	11	72
real	2	3	0	10	12	14	31	72
better	2	2	1	7	10	19	30	71
two	0	10	3	17	15	4	22	71
aware	2	0	1	6	7	31	23	70
challenges	0	22	0	3	6	17	22	70
web	0	0	0	17	11	19	23	70
thus	0	9	1	15	9	20	15	69
adaptive	0	0	0	9	7	20	32	68
concept	3	2	2	18	13	18	12	68
first	1	8	2	15	6	19	17	68
most	0	5	5	8	17	11	22	68
models	0	1	0	12	9	19	26	67
present	3	8	1	11	14	20	9	66
access	0	0	1	0	12	35	17	65
aspects	2	7	1	23	6	13	13	65
packet	0	0	0	7	23	19	16	65
than	4	16	2	8	11	13	11	65
need	2	1	2	7	12	16	24	64
only	3	7	3	14	8	15	14	64
all	4	5	3	6	10	17	17	62
environment	2	5	14	10	11	11	9	62

Word	Philosophy	Psychology	Tourism	HCI	Network (-2010)	Network (2011-16)	Network (2017-20)	In total
optimization	0	0	1	3	1	31	26	62
there	5	6	2	11	9	16	13	62
key	0	4	0	9	4	21	23	61
cost	0	0	1	3	10	24	22	60
information	4	9	1	7	7	18	14	60
loss	0	0	0	6	22	15	16	59
measurement	0	0	2	10	20	14	13	59
various	0	4	0	7	12	13	23	59
due	0	1	0	4	12	19	22	58
value	12	2	20	7	4	7	6	58
videos	0	0	0	7	0	22	28	57
during	1	12	3	12	5	9	15	57
several	1	4	0	2	14	24	12	57
about	2	7	1	13	9	12	12	56
number	0	3	1	7	3	20	22	56
computing	0	0	0	1	2	10	42	55
defined	1	3	5	5	13	15	13	55
media	0	2	0	17	6	17	13	55
methods	0	5	0	14	10	12	14	55
i.e.	0	4	2	5	7	15	21	54
providing	0	4	2	4	13	19	12	54
should	0	4	6	3	13	16	12	54
significant	0	10	6	12	2	7	17	54
transmission	0	0	0	5	8	21	20	54
understanding	0	3	1	19	3	17	10	53
what	7	11	3	6	10	9	7	53
characteristics	0	4	1	10	6	11	20	52
conditions	5	3	1	7	5	14	17	52
IPTV	0	0	0	2	46	3	1	52
monitoring	0	0	1	8	8	14	21	52
process	1	15	1	9	7	10	9	52
therefore	1	1	1	3	9	17	20	52
architecture	0	0	0	3	6	23	19	51
current	0	4	0	10	5	16	16	51
improving	0	0	2	5	6	18	20	51
optimal	0	8	0	4	3	21	15	51
communication	0	0	0	6	7	14	23	50
driven	0	1	0	3	4	21	21	50
resources	0	4	0	2	7	15	22	50
students	0	48	2	0	0	0	0	50

Appendix 4. Objective methods

Most quality of experience studies are based on subjective opinions instead of objective measurements – one may even argue that quality of experience cannot be studied without subjective opinions. The most promising option to measure experiences objectively is to apply physiological measurements. Nevertheless, many papers claim to perform objective measurements by technical means other than physiological methods. For instance, Juluri et al. [158] apply user engagement (e.g., playing time) as one objective QoE metric. Similarly, Jahromi et al. [150] use the number of clicks with web mapping applications as a part of their QoE model. Mok et al. [213] utilize different kinds of user activities, like mouse movements, to assess quality of experience. The hardest challenge is to combine all the measurement results into a coherent model.

Physiological measurements

As for the physiological measurements, electroencephalogram (EEG) has the longest history. After the first human EEG recordings by Hans Berger in 1924 [108], EEG and quality of experience occasionally appeared in the same papers, for instance in [167, 202], but without any significant influence on the quality of experience research. Several other physiological measurements are also applied during the last twenty years, like electrocardiography (ECG), electrodermal activity (EDA), facial electromyography (EMG), near-infrared spectroscopy (NIRS), functional magnetic resonance imaging (fMRI), and photoplethysmography (PPG). EDA has been called by various other names, most notably, galvanic skin response (GSR). Table A4.1 shows the incidence of these methods in quality of experience research.

Table A4.1: Appearance of physiological methods in quality of experience papers in Set Q3C.

	2004-2012	2013-2016	2017-2020	Notable papers
Papers in total	243	172	211	
ECG	0	1	7	[55]
EDA, GSR	3	2	5	[81, 168, 217, 260]
EEG	3	7	5	[38, 107, 350]
EMG	1	2	3	[185]
fMRI	3	0	4	[105]
NIRS	0	0	1	[87]
PPG	0	0	2	[245]
At least one	8	7	12	

Akhtar et al. [8], Arndt et al. [17], Engelke et al. [86], Grassini & Laumann [106], and Raheel et al. [245] provide valuable surveys on different physiological methods to analyze human emotions and experiences. In addition, Akhtar & Falk [7] discuss the complex relationship between QoE and physiological measurements. Regardless of these summaries and numerous mentions of physiological methods, only a few QoE papers have applied them.

The few papers using physiological measurements to evaluate quality of experience offer insightful findings. For example, when stressed, a person's sympathetic nervous system is activated, resulting in a measurable increase in electromyograms of muscles [327]. This method has been used to observe players in network games because it is unreasonable to interrupt a gaming session. Lee et al. [185] used EMG to assess the effect of delay on game performance. Their results suggest that even a short delay of 50 ms may have a discernible effect on an EMG record. Raheel et al. [245] used EEG, GSR, and PPG to assess emotions (happy, angry, sad, relaxed) during video and tactile enhanced multimedia sessions. They report high accuracy when the results of all three methods were combined. Virtual reality is another important application of physiological methods [260]. Gupta et al. [107] used EEG to measure emotions during an experiment. They calculated two indexes based on two EEG frequency subbands and assumed that the indexes can be used to assess valence and arousal. The most promising result is that valence correlates strongly with MOS, which provides a useful linkage between EEG studies and the large set of QoE studies based on MOS.

The above-mentioned studies, although somewhat limited in scope, indicate that physiological measurements provide significant results and can be used to complement subjective studies. The currently available methods, particularly those that are unnoticeable for the subject of the study, can be used to assess the valence of emotions during an experience. However, physiology-related signals are noisy, and their usage requires a significant amount of expertise [215]. Furthermore, a more challenging objective would be to objectively measure other dimensions of an experience, like how useful or meaningful the experience was.

Methods to assess perceived quality

This part of the appendix provides numerical data on different methods to assess the perceived (or perceivable) quality of audio and video streams. Table A4.2 provides data about to use of different methods in quality of experience papers. As to the performance of different methods presented in the table, see Akhtar & Falk [7], Barman et al. [30], Bouraqia et al. [42], Chen et al. [56], Chikkerur et al. [57], Garcia et al. [100], Maia et al. [194], Min et al. [206], You et al. [341], and Zerman et al. [343].

The most straightforward methods are to calculate mean squared error (MSE) or peak signal-to-noise-ratio (PSNR). However, they do not match well with perceived visual quality [321]. The most popular advanced method is the Structural Similarity index measure (SSIM), which is used to evaluate the perceived quality of digital images and videos by comparing the structural information between the original image and the compressed image [57, 321, 330]. SSIM is used in both HCI and networking fields to estimate, usually based on the results of MOS studies, the quality level required to avoid noticeable irritation.

Note also that the abbreviation VQM is used to refer to several concepts; V refers to video, Q refers to quality, whereas M can refer to measure, metric, model, or monitor. The most common version is Video Quality Metric also used in ITU-T standard [138].

Table A4.2: Methods to assess the perceived quality of audio and video streams

		Original reference	Set Q1	**Set Q3C**	- 2005	2006-10	2011-15	2016-20	Net	HCI	Other fields
ANOVA	analysis of variance	Fisher (1921) [93]	105	**46**	13	11	14	8	19	8	19
BLIINDS-II	BLind Image Integrity Notator using DCT Statistics	Saad et al. (2012) [255]	4	**4**	0	0	1	3	2	2	0
BRISQUE	blind/referenceless image spatial quality evaluator	Mittal et al. (2012) [209]	9	**8**	0	0	1	7	4	4	0
DIIVINE	Distortion Identifica-tion-based Image Verity and INtegrity Evaluation	Moorthy & Bovik (2011) [218]	5	**4**	0	0	1	3	2	2	0
DVQ	Digital Video Quality	Watson et al. (2001) [325]	5	**3**	0	0	2	1	2	1	0
FSIM	Feature Similarity Index	Zhang et al. (2011) [347]	8	**5**	0	0	2	3	2	3	0
GMSD	Gradient Magnitude Similarity Deviation	Xue et al. (2013) [339]	6	**5**	0	0	1	4	3	2	0
MOVIE	Motion-based Video Integrity Evaluation	Seshadrinathan & Bovik (2009) [265]	8	**5**	0	0	3	2	3	2	0
MPQM	Moving Picture Quality Metrics	van den Branden Lambrecht & Verscheure (1996) [311]	18	**12**	0	8	3	1	10	2	0
MSE	mean square error		81	**54**	0	13	21	20	43	10	1
MSSIM	Mean Structural Similarity	Kandadai et al. (2008) [166]	4	**2**	0	0	2	0	0	2	0
MS-SSIM	Multi-Scale Structural Similarity Index	Wang et al. (2003) [323]	34	**25**	0	0	12	13	15	10	0

		Original reference	Set Q1	**Set Q3C**	- 2005	2006-10	2011-15	2016-20	Net	HCI	Other fields
NIQE	Natural Image Quality Evaluator	Mittal et al. (2012) [210]	12	**8**	0	0	0	8	6	2	0
NVFM	Normalization Video Fidelity Metric	van den Branden Lambrecht (1996) [310]	4	**2**	0	1	1	0	2	0	0
PDM	Perceptual Distortion Metric	Winkler (1998) [329]	3	**1**	0	1	0	0	1	0	0
PEAQ	Perceptual Evaluation of Audio Quality	ITU-R (2001) [133]	7	**2**	0	1	1	0	2	0	0
PESQ	Perceptual Evaluation of Speech Quality	ITU-T (2001) [139]	56	**34**	0	18	10	6	33	1	0
PEVQ	Perceptual Evaluation of Video Quality	ITU-T (2008) [141]	13	**8**	0	0	4	4	5	3	0
POLQA	Perceptual Objective Listening Quality Assessment	ITU-T (2011) [144]	12	**5**	0	0	2	3	4	1	0
PQSM	Perceptual Quality Significance Map	Lu et al. (2003) [190]	2	**1**	0	0	1	0	0	1	0
PSNR	peak signal-to-noise-ratio		207	**139**	2	30	60	47	119	20	0
PSQA	Pseudo-Subjective Quality Assessment	Rubino (2006) [253]	35	**21**	0	7	10	4	21	0	0
PSQM	Perceptual Speech Quality Measure	Beerends & Stemerdink (1994) [36]	9	**5**	0	3	2	0	5	0	0
PVQM	Perceptual Video Quality Measure	Hekstra et al. (2002) [119]	6	**4**	0	0	2	2	3	1	0
RMSE	root-mean-square error		67	**39**	0	4	18	17	31	8	0
RRED	Reduced Reference Entropic-Difference	Soundarajan & Bovik (2011) [285]	3	**3**	0	0	2	1	1	2	0
SNR	signal-to-noise-ratio		82	**55**	0	6	27	22	47	8	0
SSIM	Structural Similarity index measure	Wang et al. (2004) [321][87]	122	**79**	0	8	36	35	65	14	0
ST-RRED	Spatio-Temporal Reduced Reference Entropic Differencíng	Soundarajan & Bovik (2013) [286]	11	**7**	0	0	1	6	6	1	0
V-factor	-	Winkler & Mohandas (2008) [330]	5	**4**	0	3	1	0	4	0	0
VIF	Visual Information Fidelity	Sheikh & Bovik (2006) [272]	17	**12**	0	0	8	4	6	6	0
VIFp	Visual Information Fi-delity in Pixel domain	Sheikh & Bovik (2006) [272]	4	**4**	0	0	1	3	2	2	0
VisQOL	Virtual Speech Quality Objective Listener	Hines et al. (2015) [123]	4	**2**	0	0	0	2	2	0	0

[87] It is worth noting that the number of citations in Google Scholar to the original SSIM paper [321] is amazing, 36894 (14.5.2022). That number is larger than the estimated number of citations to *all HCI papers* with 3 mentions of quality of experience (25752) shown in Table A2.1 in Appendix 2 (parameter β). Wang et al. [321] use the concept *human visual system* but it does not contain terms experience, satisfaction, or feeling.

		Original reference	Set Q1	**Set Q3C**	- 2005	2006-10	2011-15	2016-20	Net	HCI	Other fields
VMAF	Video Multi-Method Assessment Fusion	Li et al. (2016) [187]	16	**9**	0	0	0	9	7	2	0
VQA	Video Quality Assessment	VQEG (2000) [316]	29	**18**	0	0	4	14	15	3	0
VQM	Video Quality M*		73	**50**	0	14	19	17	44	6	0
VQM	Video Quality Measure	Voran (1991) [317]	2	**0**	0	0	0	0	0	0	0
VQM	Video Quality Metric	ITU-T (2001) [138]	42	**31**	0	7	15	9	26	5	0
VQM	Video Quality Model	Pinson & Wolf (2004) [239]	9	**6**	0	4	0	2	5	1	0
VQM	Video Quality Monitor	AcceptTV (2015) [3]	2	**1**	0	0	1	0	1	0	0
VSNR	visual signal-to-noise ratio	Chandler & Hemami (2007) [51]	10	**8**	0	0	4	4	5	3	0
VSSIM	Video SSIM	Wang et al. (2004) [322]	4	**3**	0	0	3	0	2	1	0

Machine learning

In this brief subsection, machine learning is discussed mainly due to its great potential also in the case of modeling of quality of experience.[88] Machine learning algorithms have become popular in recent years in QoE studies. This development is understandable because of three reasons. First, many QoE studies generate an abundant amount of data to be analyzed and interpreted. Second, the relationship between different types of data (subjective and objective) is complex and depends on numerous contextual factors. Third, the expertise in the use and development of machine learning algorithms has been considerably improved over the past decade.

Bouraqia et al. [42] state that "Artificial intelligence and machine learning algorithms have been recently used to measure the QoE objectively or to improve it." Although machine learning algorithms are a useful part of a QoE analysis, it would be more accurate to assert that they are mainly used to create a link between two sets of raw data, for instance, a set of pictures of faces using a service and another set describing satisfaction on a MOS scale. In this kind of cases, machine learning may, indeed, provide invaluable assistance for constructing a workable and understandable linkage between the data sets. Another question is to what extent the results can improve our understanding about the phenomenon under study.

As to the utilization of machine learning algorithms in quality of experience research, two observations can be made:[89] First, machine learning has been used mainly in technical research. In Set Q3C, 38 papers have applied machine learning for some research purpose – 35 of those are related to networking while the remaining 3 are HCI papers. Second, machine learning has become popular during the last years: only one of the 38 ML papers in Set Q3C was published before 2010 whereas 27 papers were published from 2017 to 2020. Examples of relevant papers related to quality of experience research include:

Aggarwal et al. [4]: Machine learning is used to predict QoE for different applications.

Anwar et al. [15]: Machine learning is used to predict cybersickness in the case of virtual reality video.

[88] I used genetic algorithms to optimize connection admission control algorithms in Asynchronous Transfer Mode networks in the 1990s. More recently, I have been involved in studies predicting mobile phone popularity by Bayesian networks [169] and mobile QoE by the random forest method [43]. Based on these experiences, I would claim (concerning QoE research) that any successful application of artificial intelligence requires human insight into the overall circumstances, human needs, and the critical limitations of the model, etc. I would, therefore, prefer methods with a comprehensible structure and operational logic that enable effective cooperation between people and artificial information systems.

[89] See also Table 3 in the History section.

Balachandran et al. [22]:	Machine learning is used to predict QoE in the case of Internet video.
Boz et al. [43]:	Random forest and support vector machine are used to predict mobile QoE.
Casas et al. [50]:	Machine learning is used to predict QoE based on QoS parameters and crowdsourced feedback.
Da Hora et al. [68]:	Machine learning is used to predict MOS based on QoS parameters.
Mitra et al. [208]:	Bayesian networks are used to predict QoE.
Mushtaq et al. [221]:	Machine learning is used to predict QoE based on QoS parameters.
Nawaz et al. [224]:	Machine learning is used to analyze the influence of gender and viewing frequency on QoE.
Orsolic et al. [227]:	Machine learning is used to predict YouTube QoE based on encrypted traffic.
Porcu et el. [241]:	Machine learning is used to predict video QoE based on facial expression and gaze direction.
Xie et al. [335]:	Machine learning algorithms are applied to optimize software defined networks as regards QoE.

Appendix 5. Small MOS experiments

In this appendix, I briefly report the results of two small studies conducted at Aalto University on a master-level course. The first part, cognitive distance, shows the difficulties of constructing a set of terms with equal steps between the consecutive levels. The other part, MOS as a basis for a utility-scale, shows that even if a MOS scale could be constructed as an interval scale when the sole application is to measure cognitive distance, that does not indicate that MOS could be used as an interval scale to model human decisions or well-being. The minimum requirement is to perform a non-linear conversion from a cognitive scale to any scale that is interpreted as a utility-scale.

Cognitive distance

This section provides some interim results of a study to evaluate the cognitive distance between different terms, including the terms used in the MOS scale, excellent, good, fair, poor, and bad. The study was carried out in 2012 as a part of the Modelling human behavior course at Aalto University. The students were asked to locate 18 terms on a scale with 50 levels. The overall results are shown in Figure A5.1 together with the results from two other similar studies. The four main observations are: First, although the studies have been conducted in different decades and with different types of participants, the overall results are similar. Second, there is not much space available over the level "excellent." Third, because the difference between poor and bad is small, it is advisable to avoid using them on the same scale. Forth, there is room and even need for a term below bad. Possible choices include unacceptable, worst imaginable, and useless.[90]

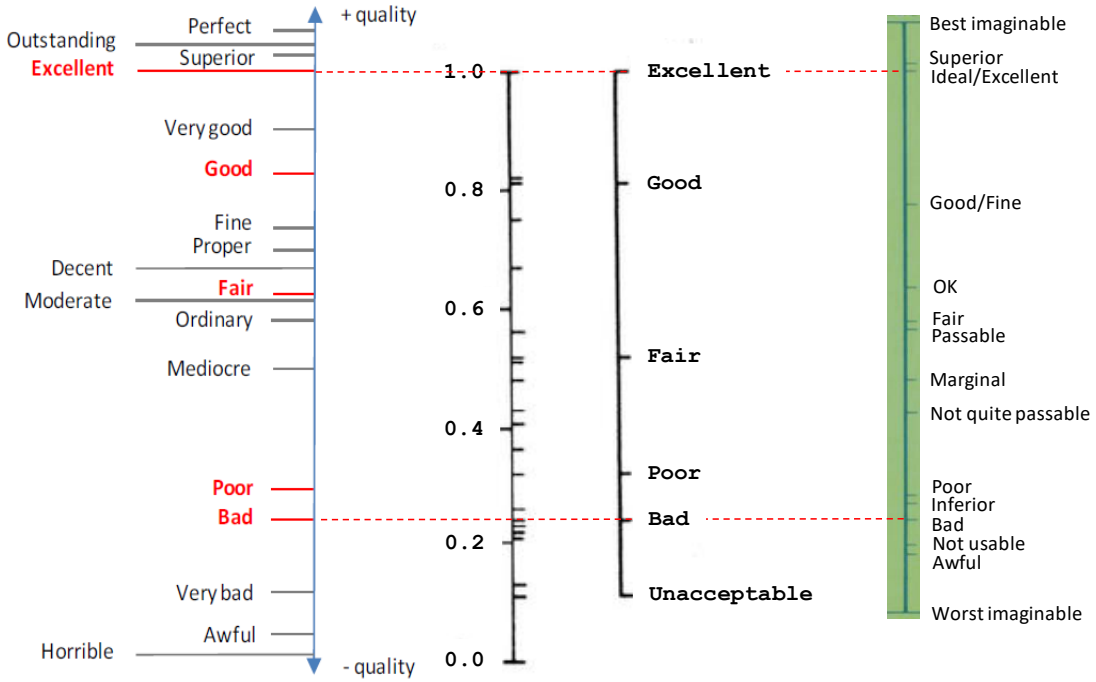

Figure A5.1: The ordering of 18 terms from perfect to horrible based on the assessment of 34 students (left), the results documented in [80, figure 1] (middle), and in [157, figure 1] (right). The scales are normalized in a way that excellent and bad are on the same level on all three scales.[91]

I am not aware of the origin of all the five terms used in the MOS scale (excellent, good, fair, poor, and bad). It might be that the set has been devised independently by different researchers for different purposes. Nonetheless, using both poor and bad on the same scale appears as a poor if not bad choice.

[90] In addition, *fair* seems to be somewhat problematic for those with a mother tongue other than English. 29 students participated in the task of assessing the set of terms, mainly from Finland and Asia. If I remember correctly, the variation in the location of fair was larger than with the other standard terms (excellent, good, poor, and bad). Unfortunately, I do not possess any more individual answers, only the summary results shown in Figure A5.1.

[91] Zielinski et al. [352] present similar scales for several languages.

There are some other approaches using different terms and labels. For instance, ITU-T [140] mentions a classification which consists of four levels: excellent service quality, intermediate service quality, poor service quality, and service not available. This set of terms appears clear and informative, but it is hardly ever used in quality of experience literature; Kerpez et al. [170] is the only exception in Set Q3C.

Another possible approach (in addition to a limited number of predefined choices) is a slider. However, regardless of the apparently linear nature of the slider, the distribution of responses is not smooth but presents patent peaks in the middle and three-quarters [232, figure 8]. Therefore, a scale with a limited number of choices seems to be more convenient, where limited can mean 5, 6, or 7.

MOS vs. utility-scale

There is no authorized or established model to convert MOS values to a utility-scale where the objective of the utility-scale is to model the behavior or well-being of a rational person. Thus, we briefly report experiments made at Aalto University to clarify the situation. The setting of the experiments was an imaginary situation in which each student made choices between two options with different qualities, for instance, between A) a phone call with quality of "fair" quality lasting 5 minutes and B) a phone call with "excellent" quality of 4 minutes and then 1 minute with "poor" quality. In all comparisons, the quality of the voice call was constant in option A while the quality in option B was first above the quality of call A for a while and then lower than in call A for the rest of the call. In both cases, the total length of the imagined call was 5 minutes. A similar experiment was organized in five courses in which 114 students made in total of 3921 comparisons (the set of options to be compared varied between the courses).

If the MOS scale is assumed to be linear, the students would prefer the call with a higher average MOS weighted by the length of the period. But that did not happen. In a more elaborate model, each MOS level is assigned a utility level. In addition to the established MOS scale from 1 to 5, two levels were added: totally useless quality has the utility of 0 per time unit while the utility of superior (or ideal) quality was located above the excellent quality. The scale is similar to that in [241], in which the lowest quality is called extremely bad, and the best quality is called ideal. The utility of excellent (5 on the MOS scale) is fixed to 100 per time unit. In addition, the model includes a parameter that defines to what extent students prefer fixed quality to changing quality. Thus, there are six free parameters: utilities for MOS values of bad, poor, fair, good, and superior, and fixed-quality preference.

The best fitting illustrated in Figure A5.2 is achieved with the following utilities: U(bad) = 39, U(poor) = 64, U(fair) = 83, U(good) = 94, U(superior) = 103, and fixed-quality preference = 9. The two pre-defined values are U(useless) = 0 and U(excellent) = 100. The unit for all the figures is utility per minute, also for the fixed quality preference.[92] Table A5.1 shows the summary of the data and the prediction of the created utility model.

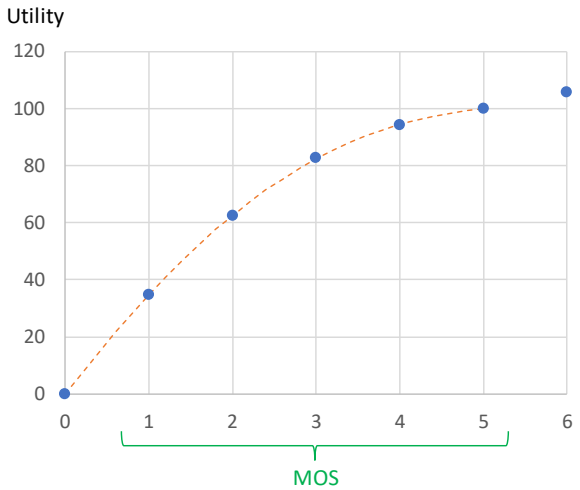

Figure A5.2: A tentative relationship between an extended MOS scale and a utility-scale. Note that zero value for utility and MOS means useless service.

[92] The fixed value preference is 36 for the period of 5 minutes, which means 7.2 per minute.

In the best case, the utility-scale might be interpreted as a ratio scale (but this issue requires further studies). Note particularly that the ordinary MOS scale does not permit a statement that quality of experience is increased by 33% if MOS is increased from fair (3) to good (4). It also is forbidden to use expressions like "X% higher QoE compared to" unless a proper ratio scale is defined and utilized. The utility-scale is more appropriate than the MOS scale in any analysis that includes several persons and their experiences. For example, a change from bad (1) to fair (3) for one person is more significant than a change from fair (3) to excellent (5) for another person. Therefore, every system optimization should be carried out on a scale that has similar properties to the utility-scale in Figure A5.2.

It should be stressed that because the results are based purely on imaginary assessments, more systematic research is necessary. I assume that the utility conversion designed based on this kind of study describes primarily the experienced utility whereas the conversion from MOS to decision utility likely require a different kind of experiment. It is possible that the research arrangement and the form of the questions described above lead to a situation in which the respondents form a kind of mixture of cognitive distance and experienced utility. If that is the case, then the “correct” conversion from MOS to experienced utility is more concave than what is presented in Figure A5.2.

Table A5.1: Results of DRM experiments. Each row represents a (mental) comparison between two options: 1) a five-minute phone call with fixed quality marked by **orange** color (e.g., on the first row, fixed quality is E = excellent), 2) a phone call with variable quality, numbers refer to the length of two periods with given quality (e.g., on the first row, there is first a four-minute period with superior quality (S) and then one minute period with good quality (G). Preference columns show the number of students preferring either fixed or variable quality. The last column shows the share of preferring the option of fixed quality predicted by the model explained in the text.

6	5	4	3	2	1	0	Preference		Share of fixed pref.	
S	E	G	F	P	B	U	Fixed	Var.	Data	Model
4		1					49	47	51 %	63 %
3			2				90	6	94 %	82 %
2		3					73	23	76 %	77 %
2			3				41	0	100 %	90 %
4			1				29	67	30 %	50 %
4				1			78	18	81 %	65 %
3			2				63	33	66 %	66 %
2			3				89	7	93 %	79 %
	4		1				46	68	40 %	66 %
	4			1			103	11	90 %	78 %
	4				1		58	1	98 %	89 %
	3		2				91	23	80 %	76 %
	3			2			59	0	100 %	91 %
4					1		54	42	56 %	43 %
4						1	67	29	70 %	68 %
3				2			38	58	40 %	55 %
2				3			79	17	82 %	81 %
	4			1			25	89	22 %	40 %
	4				1		85	29	75 %	60 %
	3			2			80	34	70 %	67 %
	2			3			104	10	91 %	86 %
		4		1			46	68	40 %	56 %
		4			1		89	25	78 %	74 %
		3		2			100	14	88 %	77 %
4						1	16	80	17 %	10 %
	4				1		3	111	3 %	7 %
	4					1	22	74	23 %	17 %
	3				2		42	72	37 %	33 %
	2				3		79	35	69 %	77 %
		4			1		12	102	11 %	13 %
		4				1	28	68	29 %	29 %
		3			2		65	49	57 %	45 %
			4		1		31	83	27 %	36 %
			4			1	51	45	53 %	61 %
			3		2		80	34	70 %	69 %
	3					2	23	73	24 %	6 %
		2				3	58	38	60 %	64 %
			3			2	48	48	50 %	24 %
			2			3	35	20	64 %	78 %
				2		3	35	6	85 %	92 %

Appendix 6. Affect and arousal – a DRM analysis

The main objective of this appendix is to provide some ideas for the future quality of experience research regarding the use of the Day Reconstruction Method [162]. However, there are some considerable limitations in this study. It has been conducted as a part of a master-level course at Aalto University without controlling all relevant confounding effects. Moreover, the main goal was to support the learning objectives of the course rather than to conduct scientific experiments. Nevertheless, some of the methods and findings may be of a general interest.

The common assumption in quality of experience research is that quality of experience can be assessed on one-dimensional scale describing the desirability of the outcome of an event (e.g., video streaming session). However, sometimes two or even more dimensions can be useful for later analysis. This appendix presents some preliminary results in which DRM is used to classify episodes on two dimensions, valence (negative-positive) and arousal (passive-active). The valence dimension is essentially the same as what is typically measured by MOS. Note also that in technical field of studies, MOS is limited to the negative side of the scale while the positive side usually depends on non-technical issues, like the quality of the content, the quality of the environment, the presence of friends, etc. As to the other dimension, arousal, it would be helpful to design a commonly applicable method to assess the level of arousal in different situations and contexts.

The DRM data was gathered on six different courses from 2010 to 2015. The main demographic data is the following:

- The total number of respondents: 266
- Gender: 104 females, 161 males, 1 unknown gender
- Citizenship: 144 Finland, 32 other western countries, 88 Asian, 2 unknown origins
- Age: 219 below 30 years, 46 above 30 years, 1 unknown age

The first finding is illustrated in Figure A6.1, which shows that positive and negative emotions are reported in different ways. For positive emotions, the most frequently used strength (above 0) is three and the distribution resembles normal distribution cut between one and six. On the contrary, in the case of negative emotions, the distribution is skewed towards zero with only a few cases in which the strongest levels are reported. It seems that three is considered as a kind of ordinary level for positive emotions whereas for negative emotions the rule is: the smaller, the more normal. Because of this difference between positive and negative, the only type of (regularly reported) episode in which negative emotions dominate is waiting.

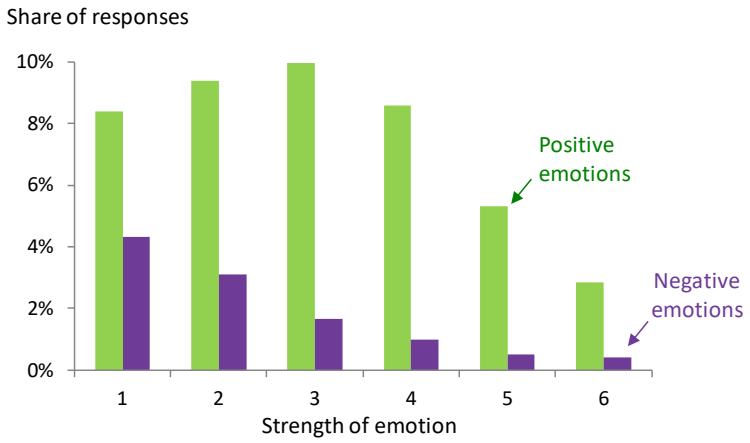

Figure A6.1: The difference between the assessment of positive and negative emotions.

There also are notable differences in the use of the scale between the participants. My original assumption (and also my instruction) was that the strength of each emotion would (and should) be in most cases relatively mild while the strongest levels should be used sparingly. Regardless of my assumption and instructions, some students used excessively values 0 and 6 and avoided the intermediate values. In contrast, some students used only the smallest values up to 2. In addition, some students reported a large number of emotions for most episodes, although the instruction was to use zero as the default level.[93] Thus, I made an effort to calibrate the results in a

[93] Students filled an Excel sheet in which an empty cell meant zero level. Thus, zero was the default choice.

way that the effect of the differences in the answering habits could be minimized. However, the effect of the calibration was so small that I decided to only remove the most skewed responses.

It would be possible, and even relatively easy, to classify all emotions on the valence scale, that is, each emotion indicates either a positive or negative feeling. It would also be possible to classify all emotions on the arousal dimension although the task is somewhat harder than on the valence dimension. However, I adopted a different approach that solely based on the DRM data instead of using my opinions about the valence and arousal of each emotion.

The approach is based on the correlation between the emotions c(i,j) shown Table A6.1. The correlation is calculated over all the assessed episodes; if two emotions are reported regularly in the same episodes (e.g., pleasure and happiness), then the correlation is highly positive. In contrast, if two emotions are reported mostly in different episodes (e.g., pleasure and boredom), then the correlation is negative. The location of each emotion (i) is defined by two parameters, $x(i)$ and $y(i)$. The logic of the analysis is that if the correlation between two emotions, i and j, is clearly positive, the emotions are located close to each other. Conversely, if the correlation between the two emotions is negative, the emotions are located far from each other.

The distance $d(i,j)$ between the locations of two the emotions i and j is simply defined as follows:

$$d(i,j) = \sqrt{\big(x(i) - x(j)\big)^2 + \big(y(i) - y(j)\big)^2}.$$

The desired distance between the two emotions is determined as $1 - c(i,j)$. Then the task is to find a solution in which the following error function is minimized:

$$E = \sum_{i,j} \big(1 - c(i,j) - d(i,j)\big)^2.$$

The result of the minimizing task is shown in Figure A6.2. The horizontal axis obviously describes the valence. The name for the vertical axis is less obvious, but because excitement and serenity were the least correlated pair among positive emotions, the vertical axis could be called arousal. In contrast, the vertical dimension on the negative side of emotions is less clear.

Table A6.1: Correlations between the reported strengths of emotions.

	Conf.	Excit.	Pleas.	Happ.	Satisf.	Love	Hope	Seren.	Bored	Frustr.	Depr.	Shame	Disg.	Envy	Fear	Anger
Confidence		0,483	0,411	0,427	0,466	0,320	0,402	0,329	0,010	0,013	0,034	0,003	0,029	0,085	0,053	0,002
Excitement	0,483		0,493	0,526	0,430	0,339	0,367	0,219	-0,117	-0,061	-0,049	0,017	-0,037	0,067	0,055	-0,040
Pleasure	0,411	0,493		0,588	0,548	0,372	0,300	0,317	-0,210	-0,202	-0,093	-0,026	-0,082	0,027	-0,026	-0,119
Happiness	0,427	0,526	0,588		0,508	0,475	0,325	0,312	-0,205	-0,202	-0,093	-0,034	-0,060	0,020	-0,030	-0,125
Satisfaction	0,466	0,430	0,548	0,508		0,359	0,343	0,333	-0,153	-0,140	-0,061	-0,023	-0,055	0,039	-0,010	-0,087
Love	0,320	0,339	0,372	0,475	0,359		0,460	0,337	-0,069	-0,068	0,016	-0,002	0,015	0,018	0,060	-0,029
Hope	0,402	0,367	0,300	0,325	0,343	0,460		0,363	0,065	0,083	0,161	0,097	0,081	0,137	0,205	0,092
Serenity	0,329	0,219	0,317	0,312	0,333	0,337	0,363		-0,019	-0,075	0,039	0,014	0,006	-0,014	0,022	-0,027
Boredom	0,010	-0,117	-0,210	-0,205	-0,153	-0,069	0,065	-0,019		0,425	0,316	0,115	0,207	0,167	0,151	0,245
Frustration	0,013	-0,061	-0,202	-0,202	-0,140	-0,068	0,083	-0,075	0,425		0,478	0,247	0,303	0,189	0,363	0,478
Depression	0,034	-0,049	-0,093	-0,093	-0,061	0,016	0,161	0,039	0,316	0,478		0,336	0,339	0,245	0,433	0,490
Shame	0,003	0,017	-0,026	-0,034	-0,023	-0,002	0,097	0,014	0,115	0,247	0,336		0,190	0,238	0,287	0,200
Disgust	0,029	-0,037	-0,082	-0,060	-0,055	0,015	0,081	0,006	0,207	0,303	0,339	0,190		0,181	0,316	0,385
Envy	0,085	0,067	0,027	0,020	0,039	0,018	0,137	-0,014	0,167	0,189	0,245	0,238	0,181		0,251	0,204
Fear	0,053	0,055	-0,026	-0,030	-0,010	0,060	0,205	0,022	0,151	0,363	0,433	0,287	0,316	0,251		0,342
Anger	0,002	-0,040	-0,119	-0,125	-0,087	-0,029	0,092	-0,027	0,245	0,478	0,490	0,200	0,385	0,204	0,342	

Pleasure	0,668	0,094
Satisfaction	0,619	-0,066
Happiness	0,607	0,233
Love	0,461	-0,252
Excitement	0,423	0,420
Serenity	0,394	-0,486
Confidence	0,306	0,267
Hope	0,187	-0,102
Envy	-0,160	-0,573
Fear	-0,305	0,238
Disgust	-0,351	0,539
Shame	-0,368	-0,462
Depression	-0,468	-0,034
Anger	-0,569	0,303
Frustration	-0,679	0,054
Boredom	-0,692	-0,296

Figure A6.2: A map of emotions based on DRM results. The size of the circle is proportional to the average strength of each emotion over all reported episodes. The formation of the emotion map is created based on the correlation data shown in Table A6.1.

Figure A6.3 presents different kinds of episodes on the constructed two-dimensional space. The results suggest that the active/passive dimension is important mainly on the positive side of the valence dimension. It seems that with negative emotions and episodes, it is not possible to distinguish a clear arousal dimension orthogonal to the strength of the negative emotions. However, the presented study is too limited to offer any strong conclusion on this issue.

According to Nawaz et al. [224] the ratings by females tend to be higher than those of males. The data in this DRM study support their finding, although the results shown in A6.4 are not statistically significant due to the relatively low number of episodes in many activity categories. Female students gave more positive evaluations for talking, social media, housework, and even sports. The only clear reverse example is waiting. Female students seem to hate waiting whereas a typical male student does not care much whether he is listening to a lecture or waiting for something or someone.

Csikszentmihalyi [65] observed that "When alone (and especially when no pressing task demands attention), the quality of experience for most people declines; depression and bad mood take over." This finding is patently supported by the results shown in Figure A6.5. Attending lectures is the only type of episode in which the reported emotions are clearly more positive alone than with friends.

In a summary, the following issues might be worth studying. First, waiting is by far the most negative type of episode for students, even worse than studying or attending a lecture. Is it possible to discern different types of waiting resulting in different levels of dissatisfaction? Second, working is more positive and more active than any kind of studying. What issues make studying so unsatisfying? Third, social media applications are not particularly rewarding activities. Why are we then spending so much time on social media? Finally, almost everything becomes more rewarding when experienced with friends. How many friends create optimal experience?

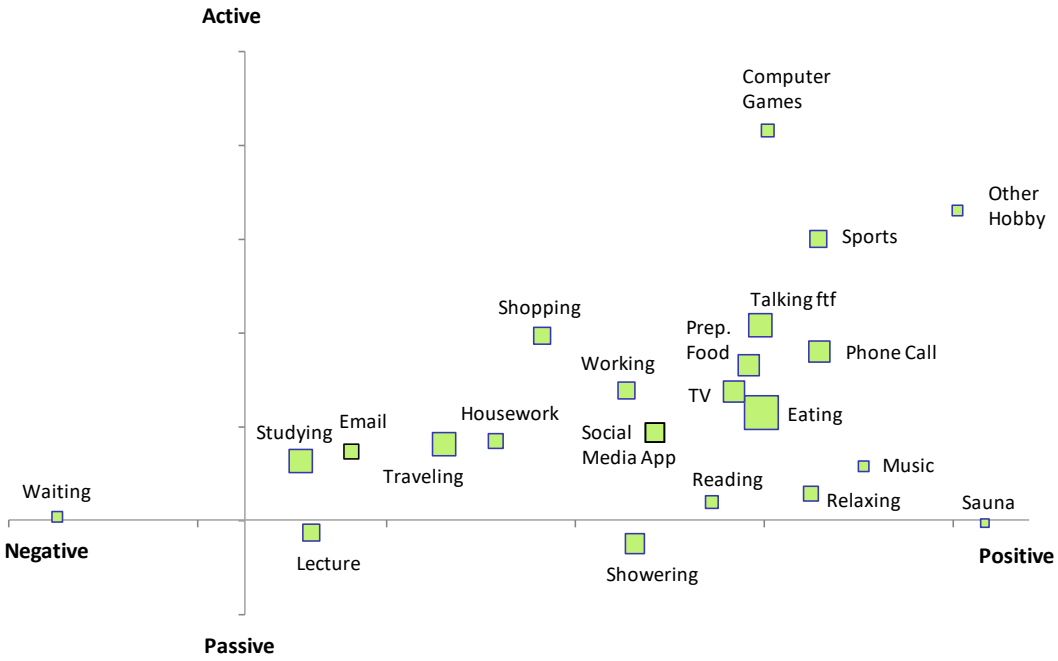

Figure A6.3: The location of episodes in a two-dimensional map. The size of the box is proportional to the number of episodes reported by the students.

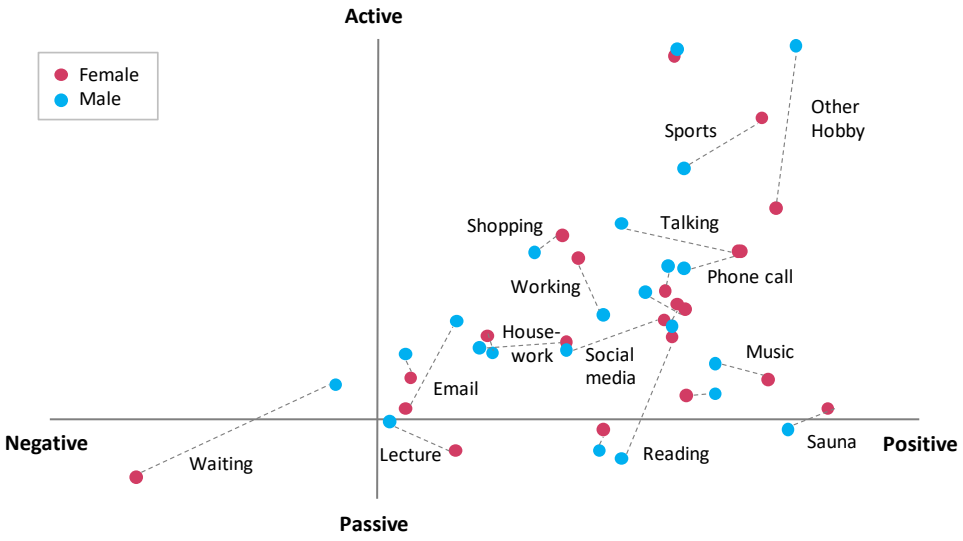

Figure A6.4: Gender difference in the evaluation of different types of episodes.

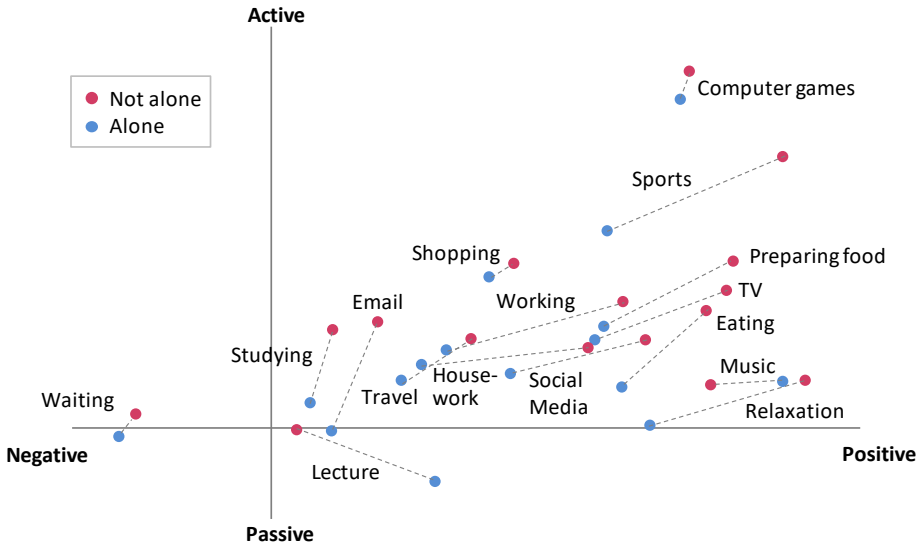

Figure A6.5: The difference between being alone and being with someone else for different types of episodes.

Appendix 7. Naming of scales

This appendix provides rumination on possible names for describing the usefulness, value, utility, or benefit of different events.

There is an incessant dilemma between the two main candidates for the purpose of life, pleasure and the well-being of society. According to one extreme opinion, promoted by the Cyrenaics in ancient Greece, the only intrinsic good is physical pleasure. At the other end of the spectrum, the only important matter is the well-being of society.[94] In modern society, we have one additional viewpoint, that of business actors and their success in gathering profits and capital. To design a generally applicable scale (that covers individual, societal, and business viewpoints), we must be cautious when selecting the name of the scale.

To assess the popularity of different terms for the scale, I studied the terminology used in the 60 articles recommended in Appendix 1. On the basis of this small study, the four main candidates are benefit, value, usefulness, and utility; all of them have been used in quality of experience literature as the name of scale. Moreover, the terms goodness and excellence have been occasionally mentioned (e.g., in [47]) mainly due to the book by Martens & Martens [198], but without any noticeable effect on the quality of experience literature. I also checked the terms worth, worthiness, desirability, and preference. Even though they appear in many articles, they are not used as the name of the scales similarly as benefit, value, usefulness, and utility are used.

Davis' infamous article [69] appears to be the main source for the term *perceived usefulness*. The definition promoted by Davis is "the degree to which a person believes that using a particular system would enhance his or her job performance." The other key term in [69] is *perceived ease of use*, whereas the paper does not address any hedonic aspect (happiness, pleasure, enjoyment, etc.) related to the use of technology. In general, usefulness often is strictly limited to the practical uses and benefits of products and services. When usefulness is used in quality of experience articles, it is typically an item in a list of factors that also contains hedonic items like the *ease of use* and *joy of use* [29, 47, 181], or *enjoyment* [276]. Usefulness is, therefore, a helpful term in the case of quality of experience analysis but not applicable as a label for a generic scale covering both practical and hedonic aspects of experiences.

Benefit is a possible term for describing the worth of an event. However, Otto & Ritchie [228] and Pilloni et al. [236] are the only articles recommended in Appendix 1 that use benefit in this meaning.[95] As to [236], a possible reason for the use of benefit is that in the paper, utility refers to electricity utility (the topic of the paper is smart home energy). In most cases, the main reason for using benefit seems to be the useful notion of cost-benefit analysis. When modeling, it would be useful to clearly distinguish between costs and benefits [298]. In the case of quality of experience, this distinction may occur on the dimension of negative and positive feelings. However, adding up costs and benefits on a feeling scale is problematic because positive feelings cannot compensate for the effect of negative feelings (see extensive discussion by Baumeister et al. [35]). Calculations are even harder if the cost is measured in money and benefits are measured in positive feelings, or cost is measured in negative feelings and benefits are measured in money. Although benefit is an appropriate and useful term in many cases, it is not an optimal label for a generic scale used in multidisciplinary research. Thus, the two main candidates as the name of the scale to assess quality of experience are value and utility.

Value is a key concept in microeconomics where the customary assumption is that rational actors make decisions based on the expected value of different offerings. Mazzucato [200] presents a praiseworthy overview of the history and philosophy of value in the context of production and economics. In contrast, experience is more about the consumption than production of services.[96] Although user satisfaction measured on a MOS scale is an indication of the value of a commercial service, it is necessary to keep in mind that quality of experience, satisfaction and value are not synonyms.

In the prospect theory developed by Kahneman &Tversky [164], the horizontal scale is named outcome while the vertical scale is named value. The two main properties of the prospect theory are the following. First, the reference point divides the value function into two parts, positive and negative, so that the psychological effect of an outcome is stronger on the negative side (e.g., losing money) than on the positive side (e.g., gaining money).

[94] Well-being may refer either to the successful future of a nation or to the combined well-being of all people living in society. The standpoint of this report is the latter.

[95] I have used the term *net benefit* when analyzing the value of communications services [240] and the *zero-benefit level* when constructing a value of time model [175].

[96] Many QoE studies address the production side of services, e.g., by attempting to optimize the implementation of video streaming services. First, an agent creates a movie, then the movie is streamed over infrastructure to a customer, and finally, the customer's brain creates the conscious experience based on the information made available by the creator and the service provider; all three agents are necessary elements for creating the experience.

Second, the relationship between outcome and value is non-linear. What is assumed to be linear is the value scale in the sense that a rational person attempts to maximize the expected value on the scale; whether people are truly rational in this sense is another issue. The assumptions made in prospect theory allow relatively simple mathematical evaluation of complex psychological phenomena.[97]

In retail and tourist domains, *experiential value* refers to customers' perceptions of products or services through direct use or indirect observation [199]. For instance, Wu et al. [332] construct a conceptual model to analyze the behavior of theme park visitors. The model incorporates experiential quality, emotional value, functional value, and experiential satisfaction. However, any wider application of their model and concepts is challenging because the concepts are integral parts of a structural equation model (SEM); if other methods are applied, the concepts do not necessarily provide clear directions.

Perceived value is a common term, especially in the tourism sector, see [23, 79, 83, 99, 276]. However, the conceptual studies on perceived value are limited and unsystematic according to Domínguez-Quintero et al. [79]. This claim is supported by the complex terminology proposed by Shin [276] containing terms like perceived utilitarian performance, and perceived usability. Ellis et al. [83] offer an interesting analysis of perceived value in their theory of structured experience. They define perceived value as "The degree of contentment an individual has with her/his investment of resources in a structured experience." The other two components of experience are affect (or valence) and delight (or arousal). This construction leads to a distinction between the quality of experience (as an emotional response) and the perceived value (as a result of a cost-benefit calculation).

Value of time is a more specific term but closely related to the quality of experience, see [148, 153, 175, 283]. The main application of the value of time has been in the evaluation of time savings as a result of better transportation and the additional cost of waiting time due to traffic congestion. A paper discussing the nature of slow travel [191] is a rare example in which the terms quality of experience and value of time are used within the same analysis. The main difference between the value of time and quality of experience models is in the fundamental nature of the scale: the value of time is measured on a monetary scale (e.g., €/h), whereas the effect of waiting time in the context of quality of experience is usually measured on an opinion score scale (see, e.g., [82, 125]). Value of time studies, nevertheless, offer valuable insight into the conversion between experiences (e.g., waiting) and behavioral and business models.

Altogether, value is a vague term used in diverse ways. The diversity is also an advantage because the term can embrace the diversity of human experiences. However, there is one problematic aspect in the use of value as a label for a generic scale, its economic connotation.[98] That is not a problem in the case of technical or business analysis, whereas, in the case of well-being analysis, economic connotation could be harmful.

The last candidate to label the generic scale is *utility*. Utility is the most widespread concept used to assess the importance of an object or event [247, 289]. An apt definition for utility is: "fitness for some purpose or worth to some end."[99] But there are different purposes and different ends, which may lead to the need to define and measure a set of utilities instead of one generic utility. Kahneman & Snell [163] distinguish three versions of utility: decision utility, experienced utility, and predicted utility. Kahneman has also made a distinction between moment-based utility and remembered utility [159]. The decision utility or the weight of an outcome in a decision model has been the main approach in modern microeconomics [165]. Experienced utility usually refers to the hedonic quality and comes close to two similar concepts, quality of experience and value of time.

Utility is regularly used in the quality of experience literature as well. For instance, [244, 250, 274] provide valuable discussions on the relationship between utility and quality of experience. Thakolsri et al. [297] maintain that utility is a function of bit rate on a MOS scale from 1 to 4.5. Phan et al. [234] define a utility-scale so that the consecutive steps in the scale of poor, fair, good, and excellent are equal, whereas the step from poor to no service is larger and leads to negative utility. The problem with their model is that the sizes of the steps between the consecutive levels (poor, fair, good, and excellent) are not necessarily equal as discussed in Appendix 5. In general, an appropriate utility-scale must the designed in a way that expected utility (that is, the sum of individual utilities multiplied by their respective probabilities) is a feasible and efficient parameter to predict customer behavior.

---

[97] The prospect theory has been adopted by several QoE researchers, see [121, 192, 299].

[98] In Merriam-Webster, the first meaning for value is the monetary worth of something, the second is related to goods and services while only the fourth refers to human values (https://www.merriam-webster.com/dictionary/value).

[99] https://www.merriam-webster.com/dictionary/utility. The most common context for fitness is biology in which fitness means the capacity of an organism to survive and transmit its genotype to reproductive offspring as compared to competing organisms (https://www.merriam-webster.com/dictionary/fitness). In the context of utility, fitness may refer to the ability to support any kind of pursuit, be it survival, well-being, successful business, or a pleasing experience when watching a video stream.

Unlike technical and business research, tourist research rarely uses the term utility, instead, satisfaction and behavioral intention are the most common variables in tourist behavior models. As a rare example, Brown [46] uses utility functions to evaluate the recreation experiences and the ensuing user behavior.

Brian Shackel [269] introduces three dimensions – utility, usability, and likeability – that together with cost define the acceptability of a product. In his framework, utility answers the question: will the product do what is needed functionally? This is a narrow approach in the sense that it seems to indicate that bad usability does affect the utility of the product. Moreover, likeability is defined by the question: will the users feel the product is suitable? The approach promoted in the report is that "Shackel's utility" is a kind of expected utility that describes how well a product with expected usability and likeability satisfies the user's needs. Then the realized utility can be determined by using an additive or multiplicative factor that describes the usability of the product and other similar factors that describes the aesthetic and other similar issues.[100]

The use of utility as a key concept allows the use of game-theoretical analysis. For instance, Chen [56] notices that "On the one hand, each participant can maximize his individual utility by choosing his control strategy; on the other hand, all or some participants can cooperate to maximize end-user QoE or total utility. Future research in either direction is promising." Whenever game theory is applied, a decision utility-scale is a necessary element to systematically describe the preferences of different agents.

Consequently, an appropriate unifying scale should satisfy the following requirements: 1) The scale must be designed to ensure equal treatment between different people regardless of their financial status because the fundamental worth of an experience is independent of the social status and wealth. 2) The scale must incorporate not only immediate pleasure or satisfaction but also meaningfulness and social aspects of the experience (see [197]). 3) The same scale should be applicable for measuring the effects of an event on the person him/herself and the effects of the event on other people and society.

As a provisional proposition, *utility* could be used as the principal name to measure and analyze the human worth of experiences, like the good/bad dimension proposed by Kahneman in [161]. Qualifiers like “experienced utility” can “decision utility” be added for more specific purposes and more sophisticated models. This type of generic scale to assess the human worth of an experience would serve as a mediator between different fields of study and between different research methods.

[100] This is the approach I adopted when I modeled the behavior of consumers using mobile services fifteen years ago. It is worth noting that expected utility and usability can also depend on the price paid by the consumer and on the marketing of the product. An early version of the complex model is explained in [240] but there is no publicly available documentation of the more recent and extensive version of the operator business model. I would call the outcome of the multiplication either decision utility or experienced utility depending on the assumptions and objective of the constructed model.