

# More than I ever wanted or just good enough? User expectations and subjective quality perception in the context of networked multimedia services

A. Sackl<sup>1,2</sup>  · R. Schatz<sup>1,2</sup> · A. Raake<sup>3</sup>

Received: 28 April 2016 / Published online: 9 February 2017  
© Springer International Publishing Switzerland 2017

**Abstract** This article addresses the overall relevance and the influence of user expectations in the context of perceptual quality evaluation of technical systems, and how information about individual expectations can be integrated into research on Quality of Experience (QoE). Based on a literature survey on expectations in the fields of psychology, service quality and consumer satisfaction theory, an existing conceptual QoE perception model is extended by explicitly including *desired* and *adequate expectations* in the quality perception process. A direct expectation assessment paradigm is proposed that uses dedicated questionnaires. Further, it is shown how expectation-related knowledge can be used to increase the accuracy of quality prediction models. In addition, experimental setups are presented that enable controlled triggering of expectations in user studies. It is discussed how the role of expectations for QoE can be investigated in future research as well as how to further validate the proposed methods.

**Keywords** Quality of experience · Expectations · User studies · Modeling

## Introduction

In general, Quality of Experience (QoE) research investigates the relationship between technical quality parameters like video frame rate, video bit rate, downlink bandwidth etc. and the subjectively perceived quality of a technical system (Möller and Raake 2014). Various non-technical influencing factors have an effect on this relationship, for example, human-factors like the socio-cultural background or demographic variables and context-factors like an individual usage vs. a usage within a social context like videoconferencing (see Reiter et al. 2014 for more information). In this article, we focus on one crucial quality-related aspect which has been hitherto largely neglected in empirical QoE-related research: the *expectations* of the perceiving subject.

Although the term “expectations” is frequently used in the context of quality perception and QoE-related research, we witness a lack of applicable concepts and methods that enable the operationalization of expectations and the utilization of related findings in empirical research. Albeit commonly used QoE frameworks and definitions highlight the importance of *expectations*, clear guidance on how to actually address this influencing factor is missing. For example, in Möller and Raake (2014), QoE relates to “... person’s evaluation of the fulfillment of his or her expectations ...”, i.e. expectations as the perceiving subject’s frame of reference. Similarly, in the Qualinet<sup>1</sup> QoE

---

✉ A. Sackl  
andreas.sackl@ait.ac.at

R. Schatz  
schatz@ait.ac.at

A. Raake  
alexander.raake@tu-ilmenau.de

<sup>1</sup> FTW Telecommunications Research Center Vienna,  
Donau-City Straße 1, 1220 Vienna, Austria

<sup>2</sup> AIT Austrian Institute of Technology, Donau-City Straße 1,  
1220 Vienna, Austria

<sup>3</sup> Department of Audiovisual Technology, Institute for Media  
Technology, Technical University Ilmenau, Ilmenau,  
Germany

<sup>1</sup> COST Action IC1003, European Cooperation in Science and  
Technology, <http://www.qualinet.eu/>, last access: 29th July 2015.

definition white paper expectations are described as key factor determining the end users perceptions and related with emotional state: “Quality of Experience QoE is the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user’s personality and current state” (Le Callet et al. 2012). Despite its primarily technical focus, the ITU-T Recommendation P.10 also highlights the relevance of user expectations as an influencing factor. Here, QoE is defined as “the overall acceptability of an application or service, as perceived subjectively by the end user [which] may be influenced by user expectations and context.” Also conceptual QoE models and frameworks like the one used in Schatz et al. (2013) explicitly include expectations as a main user-specific influencing factor.

To address the observed gap in existing research—that is, a frequent mention of the high relevance of expectations to QoE but at the same time lacking empirical and conceptual research foundations (see “[Practical Inclusion of Expectations in QoE related research](#)” for details)—this article provides a comprehensive analysis of the concept of expectations and makes it methodologically applicable to QoE research. To this end, the remainder of this paper is structured as follows: first, we provide an overview of how other domains (consumer satisfaction research, service quality research) define, operationalize and utilize user expectations in various application fields. Based on these findings, suitable existing concepts that aid in the operationalization of expectations (such as *adequate* and *desired expectations*) are used to extend an existing conceptual QoE-model. Building on this, a survey on existing QoE literature clarifies which aspects of expectations are actually relevant for QoE research. The main part of this journal paper focuses on discussing our empirical findings regarding *triggering* and *quantifying* expectations in the context of QoE assessment and further describes how to use these findings to increase the prediction accuracy of QoE models. Finally, we discuss further open research questions and required future work on expectations and QoE.

### Related work: towards a conceptual QoE/expectations model

In this chapter we present the findings of our literature survey regarding *expectations* in the context of service quality and consumer satisfaction research. The resulting findings are then used to extend existing QoE-models to finally propose a conceptual QoE/Expectations model.

Based on that, existing QoE-related literature is discussed to identify open research questions in the context of QoE and expectations.

### Expectations in socio-psychology, service quality and consumer satisfaction research

A good starting point for learning about the nature of expectations is the field of psychology in which the concept of expectations is often used to describe the processes of understanding and cognition because “in perception, we consider prior expectations” (Sternberg 2011, p. 108). For example, when someone reads a piece of written text, she has to make assumptions regarding the content, the context of the work’s creation, the author’s purpose, etc. Consequently, the “understanding at each point in the [text] was influenced by [...] existing knowledge and expectations based on [...] own experiences within a particular context” (Sternberg 2011, p. 393).

Additionally, expectations play a critical role in psychological research in the context of decision making, see Gerrig (2012, p. 243), and also in motivation and behavioral changes, see Gerrig (2012, p. 300). In the field of socio-psychology, expectations are considered as an important aspect which determines how subjects interact with others to fulfill social norms, cf. Gerrig (2012, p. 447, 453). Additionally, the well-know concept “self-fulfilling prophecy” describes how triggered expectations can lead to unforeseen output, see Gerrig (2012, p. 450).

To obtain a more practical and applicable definition of expectations, research fields like human–computer-interaction, economics, etc. yield additional informative results. The authors of Bonito et al. (1999) define expectations as “a kind of schemata that focuses interpretation processes on specific meanings and functions of communicative action” (Bonito et al. 1999). As described in Higgs et al. (2005), expectations are “pre-trial beliefs about a product or service and its performance at some future time” and expectations “form the frame of reference for satisfaction judgments” (Higgs et al. 2005). Additionally, the authors of Higgs et al. (2005) divide expectations into four main categories:

1. Forecast (or expected/predictive): user beliefs about what *will* occur in specific forthcoming action regarding a specific provider.
2. Normative (or deserved/desired): consumer perception of what *should* occur based on an assessment of what is realistic and feasible regarding a specific provider.
3. Ideal (or wished): highest level of performance attainable, independent from specific provider or brand.

4. Minimum tolerable (or adequate): minimum baseline performance acceptable, independent from specific provider or brand.

According to the authors of Higgs et al. (2005), practical implications of expectations have been investigated mostly by two different research traditions:

1. Consumer satisfaction research: here, the primary goal is to understand the user's cognitive processes, which leads to customer's satisfaction.
2. Service quality research: here, the primary object is to understand and to measure quality in service environments.<sup>2</sup>

In this context, one of the most widely used models regarding perceived service quality and expectations is the GAPS model and, building on that, the SERVQUAL model developed by Parasuraman et al. (1988). For example, in the context of information system evaluation, the authors of Watson et al. (1998) used the SERVQUAL model to measure existing expectations to compare it with gained experiences: Test users had to indicate their ideal information system and to evaluate 22 questions via a 7-point answering scale with answering options ranging from 1 = "Strongly disagree" to 7 = "Strongly agree", e.g., "The employees of these Information Systems units will understand the specific needs of their users". After that, the test participants used a particular information system implementation to perform some tasks. Then, the users had to evaluate their experiences by answering slightly rephrased questions, e.g., "Employees of this Information System understand the specific needs of its users." With this information, it is possible to calculate the expectation gap.

Similar to SERVQUAL but in the context of e-commerce, the authors of Kim et al. (2003) used the Expectation Confirmation Theory (ECT), originally developed by the authors of Kristensen et al. (1999), to measure the user-satisfaction of Web services. The satisfaction level was generated by a comparison of post-purchase evaluations of a product or service with pre-purchase expectations. Following the ECT, users generate specific expectations regarding a desired product. After a trial phase, the users form perceptions about its performance. Then the participants determine if their expectations were confirmed regarding the perceived performance. Finally, the users' satisfaction level results from the previous confirmation and the underlying expectations. In the end, a reuse or repurchase is considered or not.

In general, existing literature points out an important difference regarding asking about expectations: in the

service quality tradition the subject states expectations about what the service provider *should* offer. In contrast, in the consumer satisfaction literature the subject reveals more about his/her expectations what *will* be offered.

When it comes to QoE assessment, inquiring expectations can be problematic because for the user it is often not easy to *verbalize* expectations. For example, what does a 'fast' Internet connection actually refer to? How should end users quantify the expected quality of a video transmission? In this context, the author of Ojasalo (2001) focuses on an understudied aspect of service quality and user satisfaction, namely fuzzy, implicit and unrealistic expectations. According to the author, users have fuzzy expectations "when they expect a change but do not have a precise picture of what this change should be (Ojasalo 2001)". If these expectations are not fulfilled, for the concerned users the experienced service was unsatisfactory but they do not know why. The opposite of fuzzy expectations are precise expectations. According to the authors of Ojasalo (2001), implicit expectations are so self-evident users do not actively think about them. They only become relevant and explicit for the users when these expectations are not fulfilled. Finally, there are also unrealistic expectations which are obviously unable to be fulfilled. The authors of Ojasalo (2001) argue, that fuzzy expectations can be converted into more precise expectations via a dialog between the user and the service provider. Obviously, this qualitative approach is not appropriate in the context of quantifying expectations for QoE related research.

In the field of consumer satisfaction research, *brands* play an important role for expectations. Here, examined expectations—and their fulfillment or non-fulfillment—are based on concrete products or services, cf. Cadotte et al. (1987). In contrast, assessing the influence of brands or other marketing-related aspects is not very common in QoE research. Here, the focus is mainly on assessing the impact of the quality/performance of the technical system which is typically evaluated by test participants without having any background information, e.g., in common video studies, the responsible video content provider delivering the streamed video is considered to be irrelevant.<sup>3</sup>

According to the authors of Cadotte et al. (1987), in the context of service quality research it is not always straightforward to quantify expectations. For example, it is fairly simple to quantify the speed of the service in a restaurant by means of asking about expected seconds or minutes of waiting time. Similarly, menu variety can be

<sup>2</sup> For more information about these two traditions, please see Boulding et al. (1993), Buttler et al. (1996) and Oliver (1996).

<sup>3</sup> Most QoE researchers would agree that a broad range of influencing factors impact the perceived quality, but when it comes to empirical research or modeling attempts, softer influencing factors like brands or other marketing factors are mostly neglected.

determined by the number of food items on the menu. However, it is considerably more complicated to measure and operationalize employee-friendliness. Obviously, this particular issue can be solved via appropriate questionnaire designs. Nevertheless, quantifying individual expectations is difficult and depends on the context and the evaluated service.

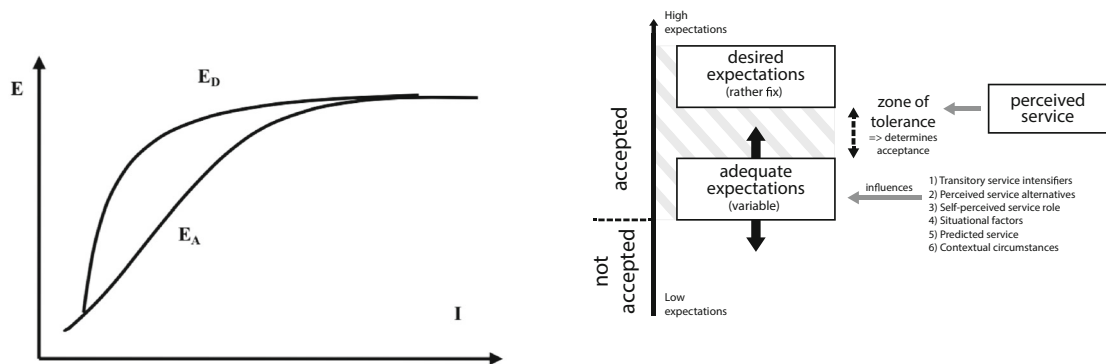
In the context of usability research, expectations can be used to identify critical usability problems. Briefly speaking, test-users in a usability experiment are asked beforehand how difficult a certain task would be, e.g., find item X on website Y. Subsequently, the users have to evaluate the difficulty of this activity after they have fulfilled this particular task. The ensuing comparison shows which design issues should be fixed immediately and which are not crucial. For example, if task A was expected by most of the users to be very easy and the ratings after the usage were mostly “very difficult”, this issue should be fixed very quickly, see Albert and Dixon (2003) for more details. Obviously, this approach is not appropriate for QoE because for the users it is often difficult to verbalize quality-related expectations, which has been discussed beforehand.

According to Zeithaml et al. (1993), there are two different types of expectations: desired and adequate expectations. The authors of Zeithaml et al. (1993) stated that the *desired expectations* are rather stable and invariant, e.g., some users are always concerned about high quality or some users are always concerned about low prices, and these basic needs do not change over time. In contrast to this, *adequate expectations* are more flexible and they are strongly influenced by the context. Between these two kinds of expectations, there is the so-called *zone of tolerance*: if the perceived service is in-between the slowly adapting desired expectations and the more variable adequate expectations, the user/customer accepts the perceived

service, see Fig. 1b for a graphical representation. In the context of typical business/customer relations, adequate expectations are influenced by (see also Fig. 1b):

1. Transitory service intensifiers: The urgency of a situation can lower the adequate expectations.
2. Perceived service alternatives: If there are alternatives available in the current situation or if it is possible to solve an issue without external support, the adequate expectations get higher, which leads to a smaller zone of tolerance.
3. Self-perceived service role: The customer tries to fulfill her role in the current process, e.g., it is not always possible to blame others for non-fulfilled expectations. Therefore, the more pretentious the level of a customer’s view on the self-perceived service role is, the higher is the level of adequate service.
4. Situational factors: These factors can lower the level of adequate service if the environmental influences are independent from a service provider. Hence, customers realize that this is not the fault of the provider and they accept a lower service level.
5. Predicted service: This is the service quality customers believe they are likely to get.
6. Contextual circumstances: For example, economic aspects are included here. A participant in the study presented in Zeithaml et al. (1993) stated that “price increases do not really drive up expectations. But my tolerance level will become more stringent/less flexible with an increase.”

There are some attempts to quantify adequate and desired user expectations in current literature. For example, the authors of Hsieh and Yuan (2009) use the Weber/Fechner-law—which originates from psycho-physics and is also used for QoE modeling—to generate a quantitative expectation measurement model which mathematically



(a) Mathematical relationship between desired and adequate expectations, originally published by [19], see also next page.

(b) Desired/adequate-expectations model. Own graphical representation of the desired and adequate expectation concept of [20].

Fig. 1 Mathematical relationship and model of desired/adequate-expectations

describes the relationship between the desired and the adequate expectations of customers regarding service providers. Figure 1a shows that desired expectations  $E_D$  are rather stable even if the stimulus magnitude  $I$  of the expectation determinants increases (e.g. personal needs, transitory service intensifiers, perceived service alternatives, customer self-perceived service role, situational factors, etc., see also Zeithaml et al. 1993). In contrast to this, adequate expectations  $E_A$  are more flexible and also increase when the stimulus magnitude  $I$  of the expectation determinants increases. Nevertheless, the authors of Hsieh and Yuan (2009) do not describe *how* to measure or quantify expectation determinants and therefore their contribution to the challenge of quantifying expectations is rather limited.

In common expectation and service quality research approaches, it is usual to capture expectations in the post-consumption phase instead of getting expectation information before a certain action is initiated, cf. Oliver (1996). Nevertheless, in the context of QoE it is more relevant to get information about users without conducting a certain experiment, so that this information can be used, for example, for MOS prediction modeling.

It is also a common assumption that test subjects have had *prior experiences* in a way that they can articulate expectations for current evaluation tasks, see Zeithaml et al. (1990). Nevertheless, some researchers assume that expectations exist even when no prior experience has been gained (see for examples McGill and Iacobucci 1992; Shirai and Meyer 1997), whereas the authors of O'Neil and Palmer (2003) state that expectations can not be generated without prior usage (Table 1).

To summarize, the nature of expectations has been examined in various research areas, e.g., human–computer-interaction, economics, psychology, etc., with the findings listed below being essential for understanding the role of expectations in the field of QoE:

- Expectations depend on a broad variety of influences, and understanding how they emerge and how they influence quality perception is not trivial.
- Expectations can have negative effects on perceived quality when being *under-fulfilled*, but also positive ones when being *over-fulfilled*. Therefore, empirical QoE research should also include test conditions in which expectations are over-fulfilled (currently, the focus of QoE research is on situations in which expectations are not fulfilled).
- In research fields dealing with service quality expectations are generally considered as measurable, e.g., by means of questionnaires. But this approach requires expectations which can be *verbalized* and which are *quantifiable*, which is not always possible in the field of QoE. For example, it might be challenging for an average end user to verbalize her expectations regarding Internet connection speed in Mbit/s.
- It is essential to distinguish between relatively stable, higher *desired expectations* and variable, *adequate expectations*, which both together influence the acceptance of a certain service.

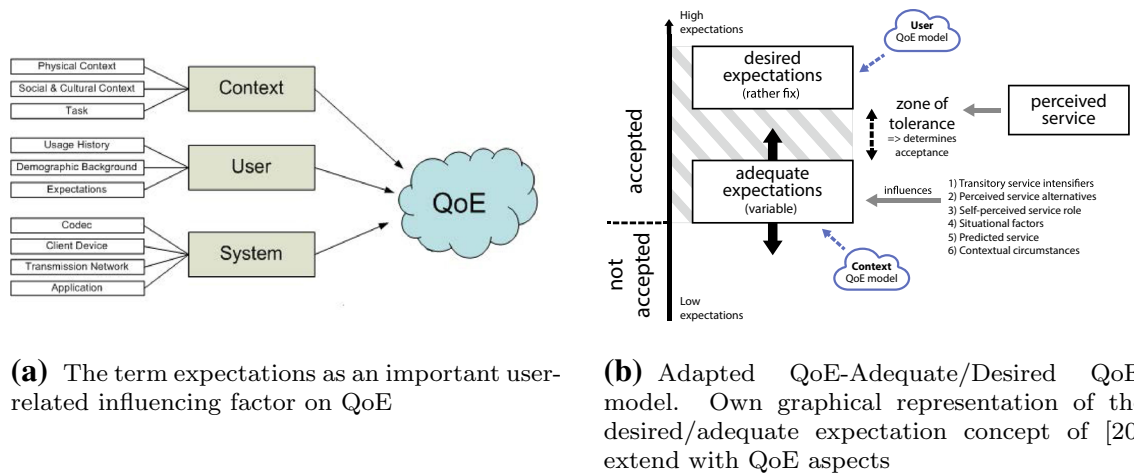
### Conceptual model QoE/expectation model

In the previous section, the current state of the art regarding expectation classification, assessment and quantification in the field of service quality and customer satisfaction research has been discussed. For the next steps it is necessary to include expectations in common conceptual QoE models in an appropriate way.

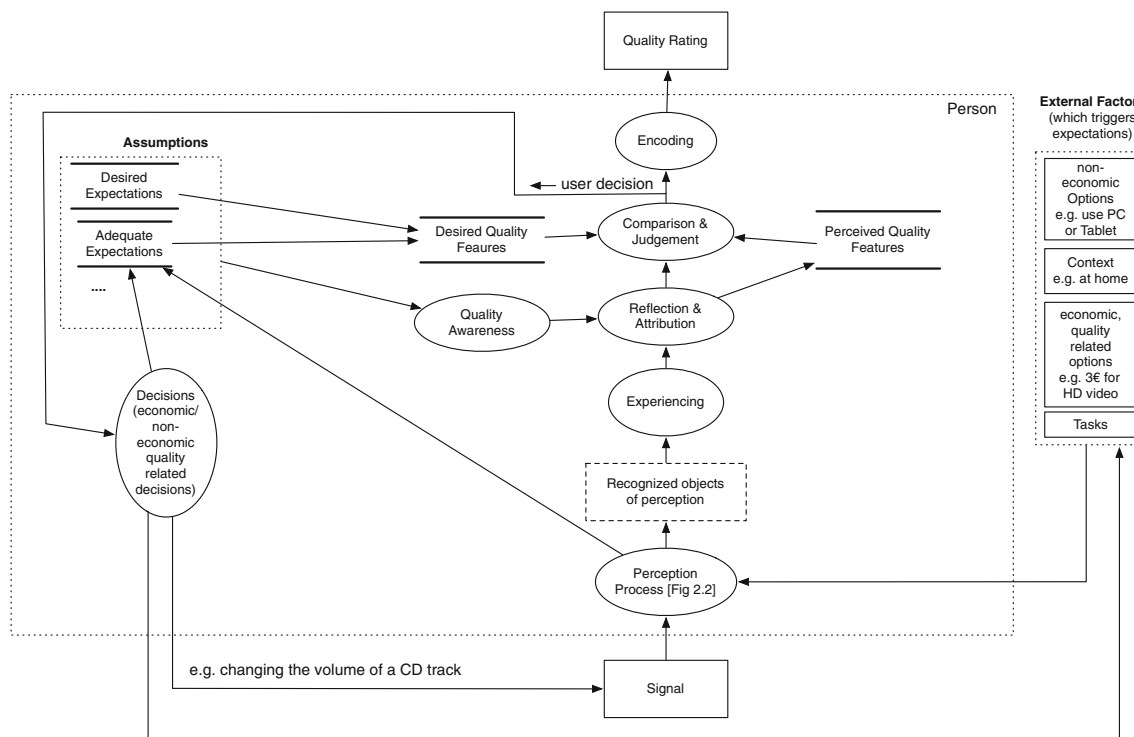
For the present work, Zeithaml's concept of *desired* and *adequate* expectations Zeithaml et al. (1993) (see also previous section) is used to integrate expectations into QoE research. As depicted in Fig. 2a, current QoE models (e.g. see Schatz et al. 2013) often consider the specific

**Table 1** Overview of the different research fields addressing expectations

Research field	Example references	Topic
Psychology	Sternberg (2011), Gerrig (2012)	Processes of understanding and cognition, decision making, motivation and behavioral changes, fulfilling of social norms
Human computer interaction	Bonito et al. (1999)	Communicative action, interpretation processes, expectations as pre-trial beliefs
Consumer satisfaction research	Rust and Oliver (1994), Boulding et al. (1993)	User's cognitive processes, tends towards brand cued expectations how is customer's satisfaction reached
Service quality research	Oliver (1996), Parasuraman et al. (1988), Zeithaml et al. (1990)	Focus on expectations and perceptions, measurement of service quality, various models e.g. SERVQUAL
Quality of experience	Menkovski et al. (2010), Monath et al. (2010)	Expectations as a relevant, but vague factor in the context of quality assessment



**Fig. 2** Generic QoE framework with influence factors and QoE/expectations model based on Zeithaml et al. (1993)



**Fig. 3** QoE-Expectations-Decision-Model. Circles represents processes, two horizontal parallel lines storages and rectangles the inputs and outputs to the user

influences of *context*-related and *user*-intrinsic variables on QoE. In terms of Zeithaml’s expectation concept, the phrase “user” may be related to the rather stable “desired expectations”, whereas the phrase “context” may be related with the context-sensitive term “adequate expectations”. Hence, a combination of the two influencing factors with Zeithaml’s model can be conceived, see Fig. 2b.

In Raake and Egger (2014, p. 23), a conceptual model of the quality-formation process—based on Jekosch (2006)

and Raake (2006)—is presented and discussed. Figure 3 depicts a (slightly simplified) representation of this model, extended by the concepts of desired and adequate expectations as part of the user-internal “Assumptions”. The perception process is triggered by signals (for example audiovisual information) and various external factors, for example contextual information representing the use case of consuming a video at home. This perception process triggers *adequate* expectations, that is, expectations that are context-specific. In perceptual terms, the *adequate* and the

more stable *desired* expectations are instantiated by the *desired quality features* used in the comparison and judgement process as the internal references that perceptual features are compared with, see Zeithaml et al. (1993). Depending on the given task, the quality formation process may lead to an explicit quality rating, e.g., resulting from a video quality questionnaire. Any kind of evaluation performed by humans may ultimately lead to a decision, based on the underlying *adequate* and *desired* expectations: For example, a Video-on-Demand content may be consumed in standard definition (SD) resolution, but for the example situation—watching a movie at home on a large TV screen—the perceived quality is too low, that is, the expectations are not fulfilled and the user decides to switch from SD to HD video content which directly influences the signal. Many different types of decisions and actions may be triggered by different signal or content-related features, such as increasing the volume, dimming the room light, or pausing or stopping the video playback.

The detailed description of the *quality formation process* can be found in Raake and Egger (2014, p. 23). In this journal article we want to give only a brief overview and explain how expectations are included in this process as depicted in Fig. 3. The “Signal” represents any audiovisual information which is perceived by the user, e.g., a consumed video, a Web site interaction etc. The fine-grained “process of perception”—which is described in detail in Raake and Egger (2014, p. 20)—results in recognized objects of perception, which furthermore influences the “experience” character of the situation, e.g., a disturbance is recognized by the user like a too long page load time during Web browsing. The so called “Quality Awareness” is triggered<sup>4</sup> which results in a “Reflection & Attribution” process, i.e., the user is aware of a quality problem, which results in “Comparison & Judgement”. During “Comparison & Judgement”, the “Perceived quality features” and the “Desired quality features” are compared which results via “Encoding” to a “Quality Rating”, e.g., the user could state “I would not accept this video quality at home”, a rating could be made on a 5 point scale etc. The “Desired quality features” highly depend on desired and both adequate expectations. We assume that desired expectations are rather stable and do not change over time.<sup>5</sup> Also “External Factors” influence the “Perception Process”:

<sup>4</sup> Additionally to disturbance situations, quality awareness can also be triggered *externally* in a laboratory setup, i.e., a researcher asks a experiment participant to evaluate the perceived quality in a particular setup, and *internally* if the user wants to buy a new TV set and compares the quality of several panels.

<sup>5</sup> Obviously, even desired expectations adopt over time, but we assume that desired expectations evolve over long-time periods like years or decades. Hence, in our model we suppose that they remain constant.

end device used (e.g. is a video consumed via Laptop or via Smartphone), situational context (e.g. is a Website accessed at home or during a train ride), is there a specific task to fulfill (e.g. booking a flight) or not (e.g. relaxing while listening to music) etc. These external factors also influence the adequate expectations—e.g. a user might be more tolerant regarding network outages during a train ride in contrast to surfing at home via a fixed-line network—which determine the “Desired quality features” for “Comparison & Judgement”. One output of the “Comparison & Judgement” could be, that the user is not satisfied with the current (quality) situation and something has to change. Hence, a “Decision” is made e.g. changing from 3G to Wifi if page load times are too high (non-economic decision) or changing from SD to HD quality during a Video on Demand Session—which might implies additional fees (economic decision). Of course, these user decisions influences the adequate expectations, e.g., switching from HD to SD video streaming lowers my adequate expectations regarding video quality, but my adequate expectations regarding interruptions are higher, i.e., I am less tolerant regarding stallings.<sup>6</sup> Additionally, some decisions also influences the “Signal”: if a user raises the Bitrate of a music streaming service, the perceived “Signal” also changes. As can be seen from Fig. 3, the decisions can be non-economic or economic ones. Economic decisions can be of short-term nature, where the user may decide to pay for getting a video in HD rather than SD, or may switch to a new service provider.

### Practical inclusion of expectations in QoE related research

In this section, we review the current state of the art in QoE-research with regard to inclusion of expectations, in the light of the model outlined in “[Conceptual Model QoE/Expectation Model](#)”. Based on the literature survey, open challenges are identified that will be empirically addressed in the subsequent sections.

#### *The relevance of expectations in QoE-related research*

Actually, many QoE researchers are aware of the existence and relevance of expectations as they form integral part of commonly used QoE-definitions (cf. “[Introduction](#)”). However, they do not explicitly assess them. In principle, when optimal quality is being achieved in a given context we can assume that at least *adequate* expectations are being met. In turn, when quality is not optimal, expectations are

<sup>6</sup> Assuming that the network connection speed remains constant, less data is transferred for SD than for HD resolution streaming. This typically results in smoother playback with less or no interruptions.

obviously not met. Hence, quality assessment can be considered as a way to indirectly assess expectations. In the conceptual model, quality results from a comparison of *desired* and *perceived* features. In principle, perceived features can be obtained from multidimensional analysis of quality, following approaches as used, for example, by Strohmeier et al. (2010) or Wältermann et al. (2010). Here, a given percept is considered to be related with a multidimensional set of perceptual (quality) features, and can be represented in a multidimensional feature space. In case of optimal quality, it can be assumed that the desired features are met by the perceived features; hence, at such an operation point, the perceived features can be seen as a measure of the desired features in that given context. However, with such indirect assessment adequate and desired expectations cannot readily be distinguished.

In the QoE community there is an increasing awareness that expectations play a key role and that expectations change and adapt over time. Nevertheless, no quantifiable results are available. For example, the authors of Monath et al. (2010) state that network providers are facing major changes in user expectations, e.g., higher awareness of the provided network quality. They also state that “an increase of usage of online services can be caused either by heavier use by existing users or an increase of the number of users. Anyway, both lead to higher expectations for performance and reliability of the services, thus increasing the demand for QoE mechanisms within the network” (Monath et al. 2010). Additionally, the authors of Micheli et al. (2013) state: “Due to the growing number of new handsets and smartphones which increases the user QoE expectations, it is important for the operators to know and to measure the UEs [User equipment] performance”. Also the authors of Mitra et al. (2011) point out, that “It can thus help in providing personalized services such as selecting a proper codec or by selecting a network interface which provides QoE based on user’s expectations.”

In the context of speech quality in telecommunication systems, the author of Möller (2000) also states that the term expectation is a rather diffuse one and it is not used in a unified way in telecommunications. He lists three components that influence expectations: the users general experience with a service, the price (e.g. more expensive often correlates with higher quality) and the nature of the connection, e.g., private call vs. informative call. In the work of Möller (2000), expectations are discussed in the context of diffusion theory (Wilkie 1994), i.e., how expectations change during the introduction and establishment of new technologies. For example, it is stated that after a new speech transmission technology has been introduced, the users start to use the new technology for other, different purposes, which leads to a decreased demand for transmission quality. While the users are

getting used to the new technology, the demand for transmission quality is increased. Also, customers can be separated into user groups, e.g., innovator or early adopter, which additionally influences the expectations regarding transmission quality during the phases of establishment.

Above survey shows that indeed there is a strong awareness and consensus in the QoE research community that expectations play a major role in end-user quality perception. However, at the same time we see very little research that explicitly address expectations in this context. More concretely, based on our survey we can identify three major challenges that research on the interplay between expectations and QoE currently faces:

1. Expectations may influence the outcome of empirical user studies, but control over these influencing factors tends to be limited.
2. Direct assessment of expectations is rather difficult and consequently, only assumptions or inferences have been made so far (e.g., on behalf of MOS ratings Cerqueira et al. 2008 or via qualitative interviews Wijnants et al. 2009).
3. It is still unclear yet how to properly extend QoE models with expectation-related influences.

In the following subsections, we discuss these three challenges.

#### *Challenge 1: Controllability of expectations in QoE-related experiments*

Although many researchers are aware of the influence of user expectations on their conducted experiments, the possibilities of actively controlling triggered expectations have been hitherto neglected: For example, the authors of Nicolas et al. (2012) investigated the difference between QoE-experiments carried out in standardized environments compared to experiments which were carried out in more realistic living room environments. The different contexts and the different expectations related to these contexts were described as the main influencing factors regarding the different quality assessment outcomes. Nevertheless, the authors did not measure, describe or quantify end-user expectations. Additionally, in Péchard et al. (2006) the authors evaluated distorted videos with different video resolutions. There, it was found that a different set of expectations seems to apply when comparing HD against SD viewing cases. Nevertheless, neither the direct influence of these expectations nor the controlled triggering of them were part of this study.

Additionally, some researchers state that expectations have been explicitly excluded in their experiments, e.g., “We remark that we do not consider other situational factors such as the users’ expectation (e.g., free vs. paid



call)” (Hohlfeld et al. 2014) and “However, from a cloud service provider’s perspective, it is challenging to gain insight into the users’ expectations and experiences” (Vandenbroucke et al. 2014). Hence, establishing controllability over evoked or triggered expectations in empirical QoE-research is highly relevant. The authors of Huong et al. (2014) developed a QoE-driven bandwidth allocation method based on user characteristics. Nevertheless, the user classification approach neglects some aspects like application, situation and also expectations regarding psychological effects. Hence, also this work demonstrates that the authors are aware of the existence of expectations and their possible impact on user experience, but specific methods to control or trigger expectations are still missing.<sup>7</sup>

So far, there are only a few expectation-related experiments in the context of telecommunication services. In Möller (2000), the author describes a user study—similar to the one discussed in Plücker (1998)—in which pairs of test participants had conversations via a portable headset and a wireline-one degraded by time-invariant impairments. For all evaluated test conditions there are no significant differences between the connection types. A separation between user groups based on the user’s previous experience with mobile phones lead to small differences between the ratings, but also here the differences are not significant. Also in Möller (2000), a study by Riedel (1998) is discussed which investigated the influence of expectations regarding making a telephone call from an Internet terminal compared to making a call from a standard wireline terminal. Similar to the previously presented study, no statistically significant tendencies regarding the impact of different expectations of connection types were found, raising the question, how expectations can be triggered in a controlled, reliable fashion.

### Challenge 2: Assessment and quantification of expectations

In the context of service quality and customer satisfaction research, the authors of Reeves and Bednar (1994) stated that customers can *articulate* how well a product or service meets their expectations. While this might be true for some kinds of products and services like cars or restaurant meals, in the context of QoE it is at least doubtful that users are always able to *articulate* expectations. For example, in the context of Web QoE users might agree that they expect a

fast Internet connection. But it might be hard for the users to actually articulate and to define what *fast* exactly means for them. In Higgs et al. (2005), the authors examine the expectations and satisfaction of an art museum with an adapted SERVQUAL-questionnaire. Obviously, it is possible to verbalize specific expectations in the context of galleries and museum experiences, e.g., range of appropriate souvenirs, employees’ willingness to help, minimize waiting lines/ticketing queues, etc. In Hartikainen et al. (2004), the authors present a study about the evaluation of a spoken dialogue system via SERVQUAL questionnaires. Five service quality dimensions were evaluated by the test participants: tangibles, reliability, responsiveness, assurance and empathy. Overall, the participants had to answer 22 items which cover these quality dimensions. Before the dialogue system was used by the participants the expectation-related part of the SERVQUAL-questionnaire was filled out, i.e., the users had to state an “accepted level” and a “desired level” regarding all items, for example “Service is fast”. After the usage of the system, a questionnaire about the perception of the system was filled out which covered the same quality dimensions. Although the discussed experiment had led to interesting insights, the authors of Hartikainen et al. (2004) critically note that the large amount of 66 questions had a negative impact on the motivation of the participants to provide correct answers. Unfortunately, no common QoE related questionnaires were used in this experiment, i.e., it is not possible to combine MOS values directly with SERVQUAL data.

Nevertheless, in the QoE research community there is rising awareness that it is generally necessary to get information about user expectations. For example, in Hirth et al. (2011) the authors state that the different quality expectations of the test participants should be considered because users with a high-speed Internet connection at home might be less tolerant regarding long page loading times compared to users with a slow Internet connection. This means that it is necessary to get information about existing user expectations to explain the quality assessment ratings obtained. Similarly, that authors of Hoßfeld and Keimel (2014) state that participants in experiments who are used to consuming video content in low resolution will rate differently than those who regularly consume video content in high resolution. Hence, their expectations are different, which makes it necessary to get information about existing user expectations e.g. via user background information in order to be able to explain quality rating behaviour.

Consequently, there have been at least a few attempts to assess quality-related expectations. For example, the authors of Yajun et al. (2014) assume that user behavior can be utilized to derive information about user expectations, i.e., according to the authors actively pausing a video

<sup>7</sup> The challenge of triggering and controlling expectations in the context of QoE is different compared to other research fields, i.e., in some research fields this kind of problem is non-existent or is perceived to be non-existent: in a field study on information system evaluation the participants of Szajna and Scamell (1993) only read descriptions of three information systems and—according to the authors—expectations for the follow-up experiments were set.

stream reflects certain user expectations and changing the video resolution while consuming a video stream also indicates user expectations. Hence, in the quoted paper, expectations are derived from user behavior. It seems that sometimes QoE and expectations are more or less interchangeable, e.g., in Cerqueira et al. (2008) the authors state that user expectations can be analyzed by measuring MOS, PSNR and structural similarity (SSIM). In this case, expectations are considered to be a result of experience and not the other way round, however, without proof.

Furthermore, expectations and the desired service features they relate to may undergo dynamic adaptation. Hence, in the context of quality assessment, such adaptation may be reflected via certain biases. For example, when a set of stimuli is presented that has a specific quality range, the usage of the quality rating scale will be different for an individual stimulus than when that stimulus is presented together with a different set of stimuli. Such effects like the range equalization bias (Zielinski et al. 2008) may partially be related to the specific focus on individual degradations or the mapping of features to an overall quality judgment.

In this respect, it has to be said that in the research described in this article, the interaction of the test paradigm with the actual study focus, namely the assessment of expectations and their role in quality evaluation, cannot completely be avoided. In Raake and Egger (2014), the term “Schrödinger’s cat problem of QoE research” was coined to describe this effect. Since the goal of this work is not to exactly quantify expectations but rather to (a) trigger them and (b) assess them in different contexts, the remaining influence of the test situation on expectations is considered to be acceptable for the studies presented in this article.

### Challenge 3: Extending QoE-models with expectations

So far, attempts to explicitly include expectations in MOS-predictive QoE-models<sup>8</sup> have been quite rare. Only the E-model defined in ITU (2005) and applied e.g. in Bacioccola et al. (2007) includes—among others—user expectations as an additional factor to calculate the resulting transmission quality rating of an audio transmission.<sup>9</sup> Equation (1) describes how the quality rating  $R$  is defined by five terms:

$$R = R_0 - I_s - I_d - I_e + A \quad (1)$$

<sup>8</sup> In the context of service quality and consumer satisfaction research it is more common to use quantified data about user expectation to predict user behavior and/or user satisfaction. For more details please see for example (Boulding et al. 1993).

<sup>9</sup> As far as we know, the E-model is the only QoS/QoE-model which explicitly includes quantified information about expectations.

According to Möller (2000),  $R_0$  describes “... the transmission rating for the basic signal to noise ratio at the virtual 0 dB point of the connection”. The terms  $I_s$ ,  $I_d$  and  $I_e$  are impairment factors representing distortions, echoes, low-bitrate codecs etc. (Möller 2000). In the original E-model<sup>10</sup> the expectations factor  $A$  stands for ‘advantage of access’. Hence, lower technical quality, for example as caused by a low-bandwidth mobile connection, tends to be compensated convenience and availability, i.e. by the fact that the user takes full advantage of being able to make a call from various locations. Therefore, the expectation factor  $A$  acts as a compensation factor for technical impairments. Nevertheless, the expectation factor  $A$  does not cover all expectation-related aspects because “... it does not take into account the special situation of the user which it pretends to model” (Möller 2000, p. 101). Hence, further effort is required to cover more expectation-related aspects in quantitative QoE models.

In the light of these aforementioned three challenges, the following research questions should be answered in order to investigate the role of expectations in QoE:

1. RQ1: Is it possible to trigger—and therefore control—expectations in empirical user QoE studies?
2. RQ2: Is it possible to assess expectations in the context of QoE in a quantitative way?
3. RQ3: Is it possible to extend quantitative QoE models with expectation information to enhance MOS prediction?

In the following sections we address these open research questions on behalf of the results of a number of empirical QoE experiments.

## Expectations in quality assessment and modeling

In this section our empirical user studies and their results regarding triggering, assessment and modeling of individual user expectations are presented. To this end, we first present the results of three QoE studies in which we actively triggered, i.e., influenced the expectations of participants. Then, we describe our method to get individual information about desired and adequate expectations via questionnaires. Finally, we explain how to utilize these information to enhance the accuracy of predictive MOS-models. Table 2 gives an overview about all studies discussed in this section.

<sup>10</sup> An extended E-Model which has been adapted for Voice over IP services is introduced and discussed in Raake (2006).

**Table 2** Study overview

ID	Research question	Year	Short description	N	Publications
TRIG 1	RQ1	2011	ADSL vs. 3G Internet access, within subjects design	49	51, 52
TRIG 2	RQ1	2013	ADSL vs. LTE Internet access, within subjects design	45	52
TRIG 3	RQ1	2011	VoD contract type, between subjects design	44	52
ASMO 1 (S4)	RQ2, RQ3	2012	Desired and adequate expectations assessment and modelling (telecom provider)	41	54
ASMO 2 (S5)	RQ2	2014	Desired expectations assessment (VoD provider)	35	–
ASMO 3 (S6)	RQ2, RQ3	2013	Adequate expectations assessment and modelling (Internet connection)	45	54

### Triggering of expectations in laboratory setups (RQ1)

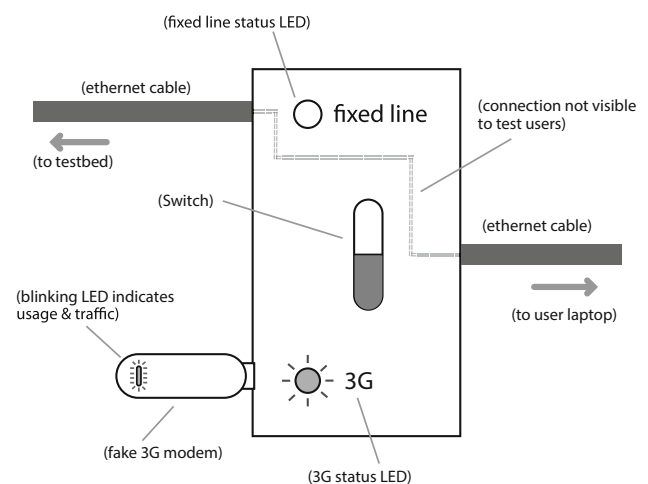
In “[Practical Inclusion of Expectations in QoE related research](#)” we stated that QoE researchers are aware of the influence of expectations on subjects’ responses in empirical user studies. Nevertheless, so far it has not been proven that explicit triggering and therefore control over an individual’s expectations is possible. Hence, in this section we present the results of three expectation triggering experiments. In general, previous publications were less comprehensive and were not embedded in a larger theoretical context compared to this journal paper.

When it comes to analyzing subjective quality assessment ratings issued in expectation trigger studies, the usual averaging approach of calculating MOS (Mean Opinion Score) values for each quality condition is not appropriate, since the individual user’s change in rating behavior (as influenced by the triggered expectations) is of interest here. Hence, we focus on the *difference* of two ratings from the same user to directly compare the influence of two different expectation triggers, e.g., wireless and wireline Internet access. In studies TRIG 1 and TRIG 2 a within-subject design was applied. In brief, *DiffRating* refers to a rating-difference by a particular user and a certain condition, whereas the average difference *DiffMOS* includes all users of a certain condition. This approach is described in detail in the following sections. In study TRIG 3 a between-subject design was applied so here the *DiffRating* approach was not possible.

#### Wireline vs. wireless internet access: study TRIG 1 and TRIG 2

In 2011 (Study TRIG 1) and 2013 (Study TRIG 2) we conducted two empirical user studies with the goal to assess how the type of Internet access (wireline vs. wireless) as *assumed* by participants and thus the different expectations triggered influence their QoE ratings. In these studies test participants had to browse several websites on a Laptop via an ADSL (2011 and 2013), via a

3G UMTS (2011) and via a LTE (2013) Internet connection. Most importantly, the users had to manually switch the connection type several times during the test procedure by themselves via a physical device called *ConnectionSwitcher*, see Figs. 4, 5 and 6. In fact, from a technical point of view both connection types *wireline* (ADSL) and *wireless* (3G/LTE) were *identical*, i.e., during the whole test the participants used a dedicated line to connect to the Internet. The *ConnectionSwitcher* was a non-functional but realistic mock-up, e.g., LEDs indicated the current connection mode and the 3G/LTE modem had built-in LEDs to indicate connection build-up and data transfer. Hence, we evaluated the *labelling effect* regarding Internet connection types. In study TRIG 1 (2011) the users browsed a custom made News Site, photos on Facebook, Google Maps and they consumed animation and music videos. In study TRIG 2 (2013) the users browsed a News Site (<http://www.cnn.com>), Google Maps, Youtube and uploaded and downloaded files. Several conditions with different maximum downlink bandwidth levels were tested, i.e., the user experienced different page load times according to the set QoS levels. After each condition—2 min. usage of certain application via wireless- or wireline-labeled connection—test users

**Fig. 4** Scheme of ConnectionSwitcher



**Fig. 5** Usage of ConnectionSwitcher



**Fig. 6** Desktop icon which indicates current connection mode

had to evaluate the subjective experience regarding network speed via a standard 5-point ACR scale, ranging from “excellent” to “bad” (ITU 2014). Overall, 49 test users participated in Study Trigger 1 (25 female and 24 male, mean age was 32 years) and 45 users participated in Study Trigger 2 (23 female and 22 male, mean age was 36.6 years).

In common QoE/quality related studies MOS (Mean Opinion Score) values are calculated. This approach is appropriate for the evaluation of (single) events e.g. evaluating browsing a Web page with 1 Mbit/s downlink bandwidth. In our studies we want to compare two ratings from a single person which might be influenced by Internet connection labelling (wireline vs. wireless). Therefore, we want to focus on the *difference* of related ratings. Therefore, *DiffRating* refers to a certain user *i* for a certain condition e.g. browsing Google Maps with a specific downlink bandwidth:

$$DiffRating_i = RatingWireless_i - RatingWireline_i \quad (2)$$

Based on *DiffRating<sub>i</sub>*, the average *DiffRating<sub>j</sub>* overall users *i* can be calculated, which is defined as *DiffMOS*:

$$DiffMOS = \frac{\sum_i DiffRating_i}{i} \quad (3)$$

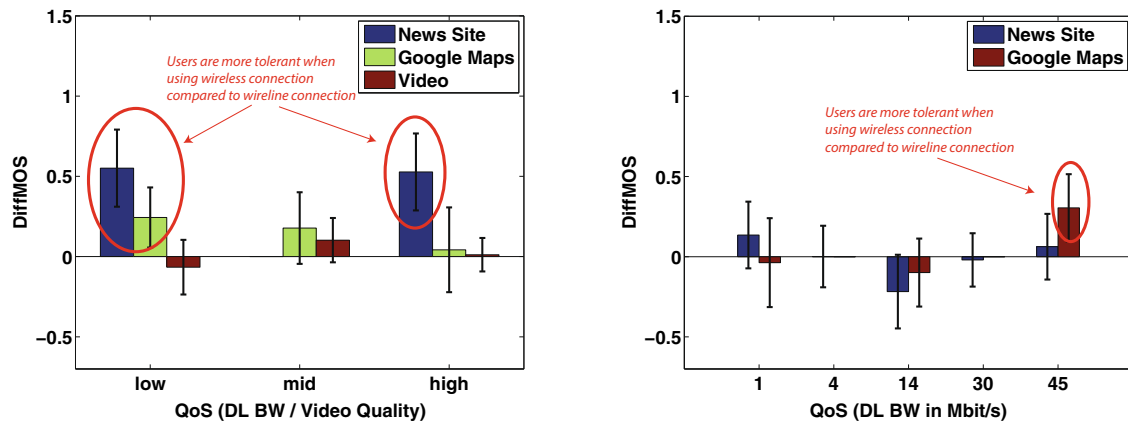
If *DiffRating<sub>i</sub>* or *DiffMOS* is different from zero, a labelling effect occurs. Nevertheless, a threshold is needed to distinguish between significant differences caused by labelling and low differences which are caused by, e.g., noise or inaccurate rating behavior. To calculate this threshold we used the 95%-confidence intervals of the *DiffMOS* values. In the CDF-plots (e.g. Fig. 8) the gray area represents the corresponding threshold, i.e., only ratings outside this area should be considered as significant. In Fig. 7, the significance of *DiffRatings* is displayed by confidence intervals.

As depicted in Fig. 7, there is a significant labelling effect in study 1 for Web usage—browsing a News Site or Google Maps via 3G or ADSL—for *low QoS* scenarios i.e. users are more tolerant if a wireless-labeled connection is used in contrast to a wireline-labeled connection (Note: both connections are technical identical!). Figure 8a shows, that 71% of the *DiffRatings* are significantly positive, i.e., wireless-labeled connections are more positively evaluated whereas only 17% of the ratings indicate an opposite effect.

For mid/high QoS situations—users experienced smaller page load times while connected via 3G or ADSL—the labelling effect is relatively weak: Only for browsing a News Site with high downlink bandwidth levels the users are more tolerant when a wireless connection is used, see Fig. 8a.

In contrast to this, in study TRIG 2 labeling study in 2013 no significant labelling effect occurs expect for browsing Google Maps and high QoS conditions, see Fig. 7b.

Hence, there is a difference regarding the Internet connection labelling-effect between the 2011 and the 2013 study. The main difference between Study TRIG 1 and TRIG 2 is the labeling of the wireless connection: In 2011, the participants thought they used a *UMTS* connection, whereas in 2013 the participants thought they used a *4G/LTE* connection. To get more information about user expectations we asked our participants about their expectations regarding Internet access via a questionnaire, the results are shown in Fig. 9: In 2011, there was a clear difference between the expectations regarding high-speed connection between fixed and mobile internet access, i.e., for wireless connections like 3G high speed access was not absolutely mandatory. In contrast to this, in 2013 there were no differences regarding the expectations of high speed access of wireline and wireless access, see blue circle/arrow in Fig. 9. Additionally, there was also a change regarding connection reliability expectations from 2011 to 2013: whereas both connection types should be very reliable in 2011, in 2013 the users did not expect the



(a) Study TRIG 1: DiffMOS results for 2011 Internet connection labeling experiment (ADSL vs. 3G) (b) Study TRIG 2: DiffMOS results for 2013 Internet connection labeling experiment (ADSL vs. LTE)

Fig. 7 MOS results regarding impact of Internet connection labeling effect

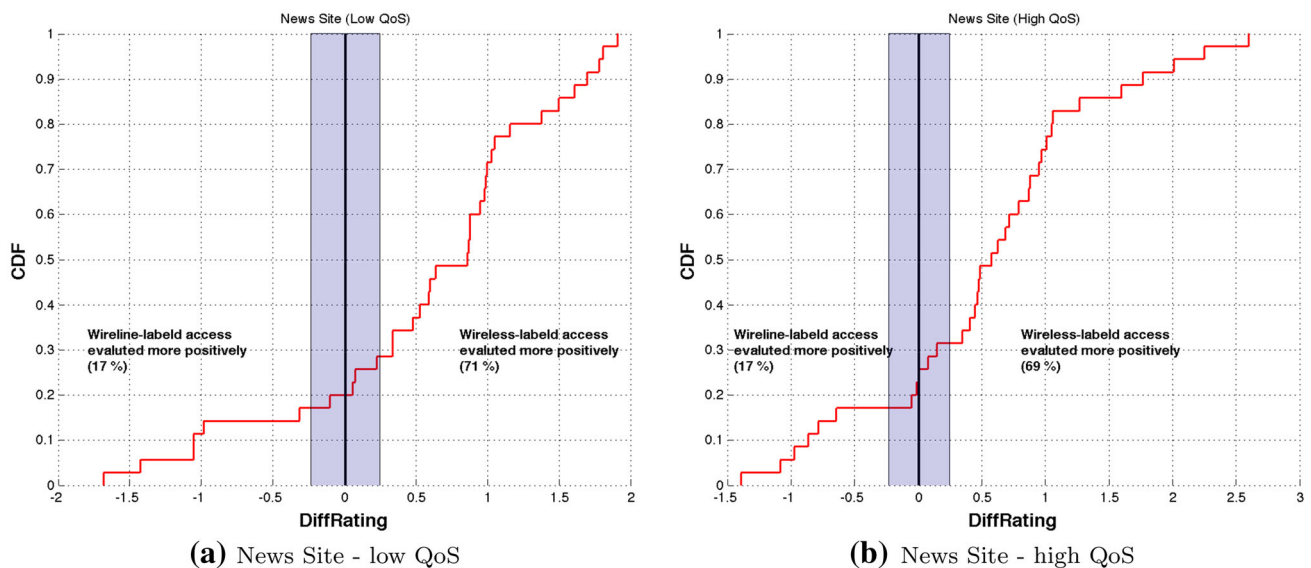


Fig. 8 Study TRIG 1: 2011 labeling study—CDFs for browsing a news site

same reliability for LTE, see red circle/arrow in Fig. 9. The emergence of expectations and their adaption over time is not handled in this paper, please see “Conclusions and Future Work”.

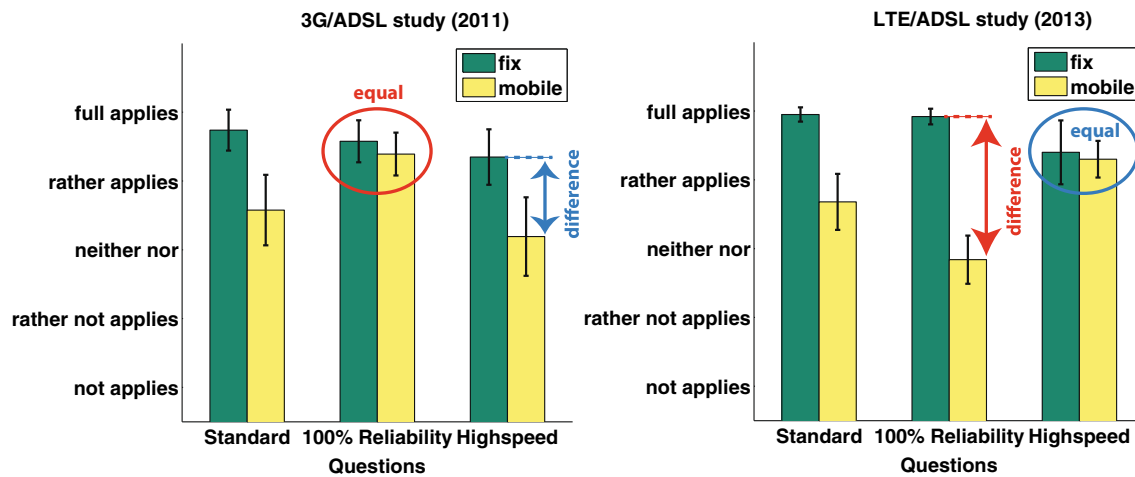
To sum up, we can verify that it is possible to reliably trigger expectations in the context of Web QoE. In our two experiments (Study TRIG 1 and TRIG 2) the labelling effect only occurs for low bandwidth settings. We also showed, that expectations change over time, nevertheless we are not able to examine the reason behind.

Video on demand contract classes: study TRIG 3

In this experiment which took place in 2011, the influence of differently priced Video-on-Demand contract on video

quality perception was evaluated, i.e., it was examined if high-priced contracts shift quality expectations. Our test participants were randomly assigned to one of three Video-on-Demand contracts (gold, silver and bronze) differing in available movies, support levels, placement of commercials, and mainly the hypothetical price charged to the user, see Fig. 10. The three contract types were presented to the users on a large TV screen and afterwards every user was assigned to one contract with the hint to use this contract type while watching the following video snippets. Hence, the participants had to imagine to use this contract while a video was consumed.

After the participants were assigned to a VoD contract, each test user watches three short video clips from the genres Action, Documentation and Sport in three different



**Fig. 9** Results from the assessment of general expectations w.r.t. fixed and mobile internet access. First question: “Mobile/fixed internet access is standard.” Second question: “I expect 100% stability from my mobile/fixed internet access.” Third Question: “Mobile/fixed high-speed internet access is very important”



**Fig. 10** Description of the different Video-on-Demand contracts presented to participants (color figure online)

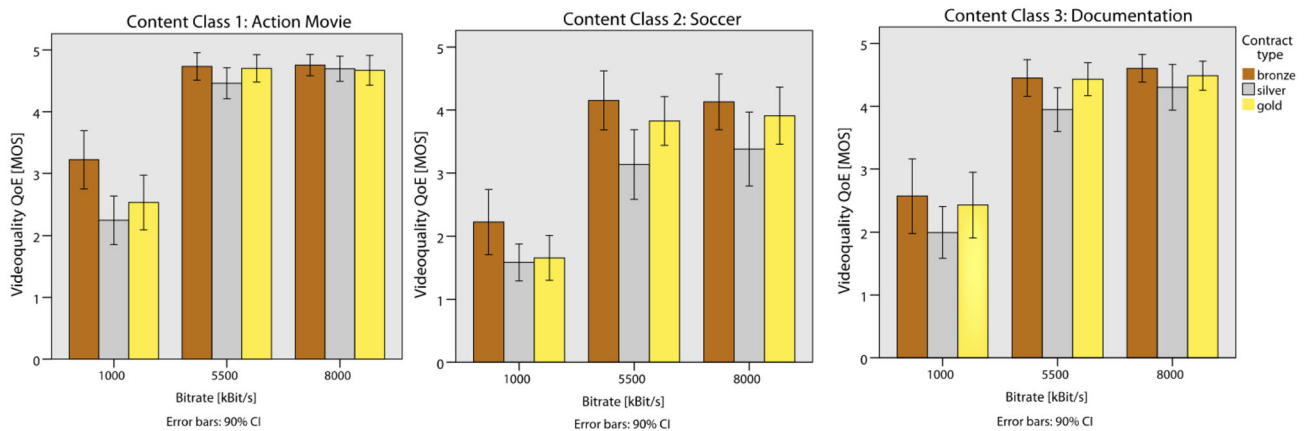
quality levels on a flat screen television (h.264 encoded 1080p/i videos with average bitrates of 1000, 5500 and 8000 kBit/s). After each video clip was presented in a particular quality level, the user had to immediately rate the video quality using an ACR video quality evaluation scale.

Overall, 44 users (22 male and 22 female) participated in our study. The mean age was 36.8 years. Classified in age groups, approximately 40% of our users were between 18 and 29 years old, 32% were between 30 and 44 years old and 28% were older than 45. Most of our users were employed (48%) while 28% were students. More than 93% of the test users were familiar with YouTube, more than 75% of them used this service once a week or more frequently. Most of the YouTube users consumed music

videos (67%) while movies and fun videos are not consumed very intensively (20%).

Only 5 users have had experiences with video on demand platforms (2x iTunes, 2x A1 Videostore, 1x UPC on demand; Reminder: study took place in 2011). Whereas test users have not watched music videos, documentations or animation movies before, one user has watched TV shows. Each month, they have spent between 2,5 and 9,9 on such services resulting in a mean of 5,48.

Figure 11 depicts the rating results for all three content classes (action, soccer, documentation). For higher video bitrates (1000 and 5500 kBit/s) there were no clear differences regarding the assigned VoD contract. There is a small, but not significant tendency, that users with a silver



**Fig. 11** VoD MOS rating results for different contract types and quality levels

contract evaluate the presented video quality more critical than users with gold or bronze contract. It seems that the provoking attempt (Imagination of using the VoD-contract while evaluating the presented video quality) might not be severe enough. Instead of imagining a certain contract, a more realistic approach should be used which could be for example a field study with a real VoD service and different contract types.

#### Conclusions about triggering

In this section we have demonstrated that triggering expectations is a complex task and not all experimental test setups lead to satisfying results. Hence, we recommend to utilize sophisticated triggering mechanisms (i.e. highly realistic emulation of contexts or services) like the ConnectionSwitcher used in studies 1 and 2. Beside its realistic functionality (e.g. blinking LEDs), the context of the switch was plausible, i.e., accessing the Internet via different connection technologies. In contrast to this and according to our results, asking participants to just *imagine* a certain situation—like using a certain VoD contract like in study 3—is not sufficient to trigger expectations in a valid way and thus should be avoided.

#### Assessment of quality expectations (RQ2)

In “Expectations in Socio-Psychology, Service Quality and Consumer Satisfaction Research” we discussed the concept of desired and adequate expectations (Zeithaml et al. 1993). In the following section we describe our approach to *quantify* individual expectations regarding quality, based on the differentiation between desired and adequate expectations. As already stated in “Related Work: Towards a Conceptual QoE/Expectations Model” quantifying quality expectations is a significant task. In this regard, our method is a first approach to address this challenge.

Evaluation of expectations has been done in QoE related research before, e.g., in a study discussed in Vandembroucke et al. (2014) the users who participated in a QoE experiment had to describe their Dropbox session experience after usage (e.g., Much worse than I expected or Much better than I expected). Nevertheless, our approach tries to get information about adequate expectations *before* any specific test condition or task has been executed.<sup>11</sup> Hence, a questionnaire design is needed which supports this research approach. In general our users fill out these questionnaires before a particular evaluation task was conducted. The questionnaires can be found in the annex of this paper.

#### Assessment of desired expectations

According to the authors of Zeithaml et al. (1993) desired expectations are rather stable and are fairly independent from context. For example, some users are *generally* more economy-driven than quality-driven, i.e., for them it is more important to save money on contrast to spend more money on exchange for higher technical quality. On the other side, there are generous users who *generally* prefer higher technical quality which is of course more expensive. If they were asked directly, users would clearly state that both aspects (quality and price) are relevant for them, hence, direct asking might not be the best approach to obtain information about desired quality-related expectations. One way of indirect questioning is using *ranking questions*, which are common in other research fields, e.g., consumer research, see Munson and McIntyre (1979). The concept of indirect asking was firstly applied by Rokeach (1968) who examined the importance of individual values. In his surveys

<sup>11</sup> Additionally, the question about expectations-fulfilling influences the following quality rating with standard ACR-scale. So, we decided to avoid direct asking of expectations-fulfilling during the test.

participants had to arrange 18 values (true friendship, mature love, self-respect, etc.) into an order of individual importance to them. Obviously, not all worthwhile values can be evaluated as most important, hence, a trade-off is needed.

To get the information about desired expectations, we asked our test participants in study ASMO 1 and ASMO 2 (see Table 2) to *rank features* of service providers, i.e., telecommunication providers or Video on Demand (VoD) vendors.<sup>12</sup> In our Web QoE study ASMO 1, the test users had to rank preferable features regarding their individual importance: high network speed, low monthly fees, short contract commitment, good support via E-Mail and telephone and unlimited download volume. The ranking of the item “high network speed” has been used as proxy for gauging the individual desired expectations regarding the desired, technical quality of an Internet connection at home. For example, a generous, quality-aware person would rank the item “high network speed” on top (=1) and the item “low monthly fee” somewhere below e.g. rank 3 or 4. In opposite to that, a thrifty person would rank the item “low monthly fee” on 1. In our Video QoE study ASMO 2, we gathered information about desired expectation from our users about VoD providers. The rank items were: large amount of available movies, excellent support, low costs, excellent video quality and short contract duration. Overall, 41 users participated in study ASMO 1 (22 female and 19 male, mean age was 36.6 years) and 43 test users participated in study ASMO 2 (31 female and 12 male, mean age was 33.2 years).

Tables 3 and 4 show the resulting ranking distributions: For desired expectations in the field of telecommunication providers, the item “high network speed” and “low monthly fee” are similar important for the user (position 1 and 2). High network speed is mostly ranked on 1st and 2nd positions, whereas low monthly fee is ranked mostly on 1st and 3rd position. In contrast to this, desired expectations in the field of Video on Demand Service providers are different: low costs (position = 1) are more important than video quality (position = 3). Figure 12 shows the histograms of the differences between desired costs and quality for both studies ASMO 1 and ASMO 2. Interestingly, regarding the difference distribution for study 5 (Fig. 13b) for many users the quality is more important than the costs. Hence, to get valid results regarding desired expectations, the calculation of the avg. rank is not sufficient, also the distribution of the differences must be considered.

<sup>12</sup> It was not possible to rank two items at equal positions.

### Assessment of adequate expectations

According to the authors of Zeithaml et al. (1993) adequate expectations are—in contrast to the more stable desired expectations—more flexible and are influenced by the context. To quantify adequate expectations, we asked our test participants in study ASMO 3 several questions which included specific details. For example, in the context of Web QoE we asked about specific tasks (browsing a News Site) and a specific context (accessing the Internet at home). For all questions there were 5 answering options, e.g., for the question regarding the download duration there were the five answering options 10 s, 30 s, 1 min, 1 min 30 s and 2 min. For the sake of comparability, for the results calculation presented in Table 5 the position of the item was used—1, 2, 3, 4 and 5—instead of dedicated values e.g. the durations in seconds or minutes. Overall, 45 users participated in Study Trigger 2 (23 female and 22 male, mean age was 36.6 years)

- “How fast should be your {homemobile} Internet access be when you browse a News Site (Answer: Mbit/s)”
- “How fast should be your {homemobile} Internet access be when you download a 50 MByte file (Answer: Mbit/s)”
- “How fast should a Website be loaded at {homemobile} via your Internet access (Answer: seconds)”
- “How long should it take to download a 50 MByte file when you use your home/mobile Internet access (Answer:seconds)”.

These resulting eight questions were accompanied by questions regarding the answer difficulty (5 answer options ranging from “very easy to answer” to “very difficult to answer [I could not answer it]”). Hence, each test participant had to answer 16 questions. Table 5 depicts the resulting mean and SD values. When asked about *durations*—e.g. “How long should it take to download a 50 MB file”—the users stated that these questions were easier to answer than questions about specific technical quality features, e.g., downlink throughput in MBit/s. Hence, to get information about adequate expectations from users regarding a certain situation (e.g. downloading a 50 MB file at home) it seems that it is more expedient to ask about directly visible quality features like waiting/downloading time.

### Conclusions about expectation assessment

In this section we have demonstrated how to assess desired and adequate expectations via dedicated questionnaires. In order to show the merit of our approach, we gathered results for a range of services—instead of focusing on a

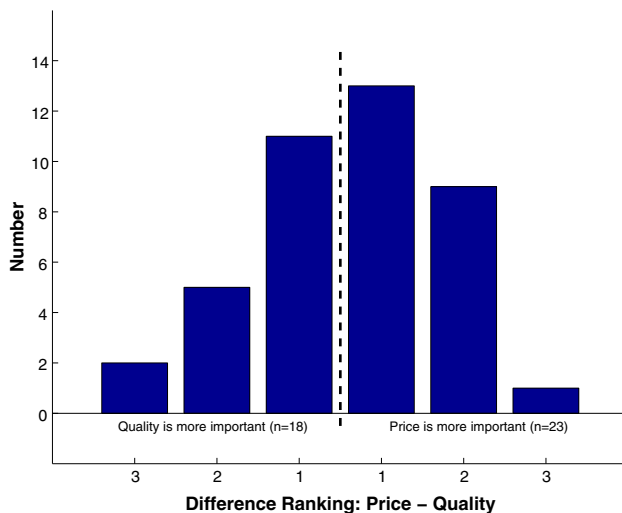
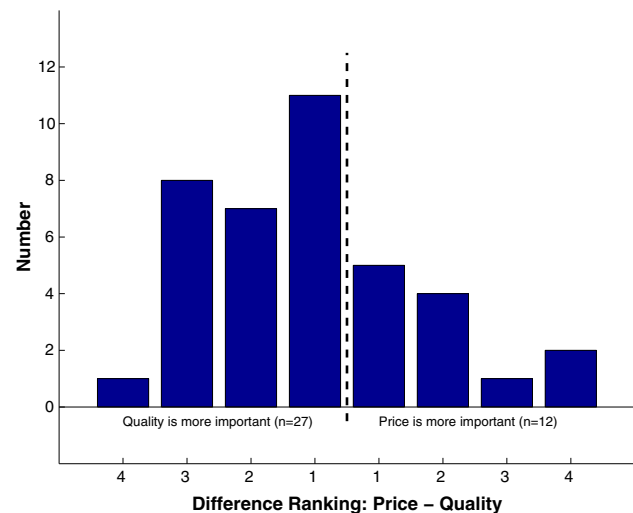


**Table 3** Study ASMO 1: desired expectations regarding internet service providers

Item	Ranked position					Ø Rank	Pos.
	1st	2nd	3rd	4th	5th		
High network speed	17	17	5	2	0	1.80	1
Low monthly fee	18	9	11	3	0	1.98	2
Short contract commitment	0	4	7	18	12	3.92	4
Good support via e-mail and telephone	0	1	6	13	21	4.31	5
Unlimited download volume	6	10	12	5	8	2.98	3

**Table 4** Study ASMO 2: desired expectations regarding video on demand providers

Item	Ranked position					Ø Rank	Pos.
	1st	2nd	3rd	4th	5th		
High video quality	12	3	9	12	3	2.77	3
Low costs	14	16	5	2	2	2.03	1
Short contract commitment	2	9	11	10	7	3.28	4
Good support via e-mail and telephone	1	0	4	11	23	4.41	5
Large amount of videos	10	11	11	3	4	2.49	2

**(a)** Study ASMO 1: Difference between price and quality desired expectations**(b)** Study ASMO 2: Difference between price and quality desired expectations**Fig. 12** Differences between quality and price

single topic—e.g. VoD provider, telecommunication provider, etc. In the following section, this information is used to enhance the prediction accuracy of quantitative QoS/QoE-models.

### Extending quantitative QoE models with information about expectations (RQ3)

In “[Assessment of Quality Expectations \(RQ2\)](#)” we have demonstrated how to assess desired and adequate expectations, in this chapter we demonstrate how this information can be used to extend QoE models in order to enhance

prediction accuracy. Table 6 provides an overview of the involved user studies.

In general, there are several ways to generate quantitative QoS/QoE-models, e.g., via machine learning techniques like decision trees (Menkovski et al. 2009) or neuronal networks (Aguiar et al. 2012). It is also common to gain less complex solutions to describe the relationship between technical and perceived quality, e.g., curve fitting like discussed in Sackl et al. (2013). In Sackl et al. (2013) the relationship between various initial delay lengths in music and video streaming scenarios and the perception of these delays (e.g. how annoying was it for the user) is

**Table 5** Overview about questions regarding adequate expectations: mean and SD

	Fixed Internet access at home				Mobile Internet access			
	Expectation		Difficulty		Expectation		Difficulty	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
<i>Site</i>								
Dur.	1.48	0.63	<b>1.61</b>	0.79	2.09	0.74	<b>1.93</b>	0.97
TP	3.98	1.52	2.70	1.53	2.55	1.50	2.89	1.35
<i>DL</i>								
Dur.	2.16	0.94	2.05	1.08	2.73	0.95	2.10	0.94
TP	2.82	1.78	2.82	1.50	3.61	1.62	2.80	1.50

Bold ones indicates avg. easiest question. Site = Browsing a Website, DL = Download of a 50 MB file, Dur. = Duration of download/page load in seconds. TP = Throughput in MBit/s

modeled via logarithmic relation. Whereas in Egger et al. (2012) the authors demonstrate that exponential functions in general are an appropriate way to model the relation between bandwidth and MOS for Web applications. However, both approaches—fitting and machine learning—have their justifications but we want to demonstrate how quantified information can be included in QoE-models and so the transparent method of fitting is more adequate than a machine learning based, black box approach.

Hence, in the following subsections an exponential fitting approach for the specific user studies are presented and evaluated. Subsequently, individual information about desired and/or adequate expectations is added and the extended model is evaluated to determine if this additional information enhances the MOS-prediction of the model.

### Modelling of desired expectations

In Study ASMO 1 (see Table 6), our test participants had to browse Google Maps (Satellite View, see Fig. 14a) via three different downlink bandwidth values: 256, 1024, 4096 kBit/s. After the tasks were completed, the users filled out a questionnaire to get their quality impression regarding the perceived speed of the internet connection (“How do you perceive the speed of the Internet connection?” with answer options ranging from 1 = bad to 5 = excellent). Additionally, we gathered information about desired expectations with ranking questions (see “Assessment of Desired Expectations” for more details).

**Table 6** Overview of studies for modeling

ID	Type of expectations	Objective
Study ASMO 1	Desired	Web QoE (Google Maps)
Study ASMO 3	Adequate	Web QoE (File Download)
Study ASMO 1	Adequate and Desired	Web QoE (News Site)

Finally, we use the position of the ranked element “Importance of Network Speed”, which value was between 1 (generally very important) and 5 (generally not important) to determine the desired expectation.

Figure 13a depicts the individual quality ratings (green dots), the resulting MOS values (blue bars) and the resulting exponential fitting curve. The first line of Table 7 shows the pure QoS/QoE model with the resulting adjusted  $R^2$  value (0 = no fitting of the model with the underlying data; 1 = perfect fitting) and the root-mean-square deviation.

Please note regarding this study and the following studies: the fitting curves are based on the *individual ratings*, which is in contrast to common QoS/QoE-fitting approaches which base on the resulting MOS—i.e. a single average value over all ratings—values. Hence, the resulting  $R^2$ -values (based on many individual ratings) are lower compared to  $R^2$ -values in other studies which are mostly based on only a few MOS-values.

Next, the pure QoS/QoE-model was extended with the additional additive, linear factor  $exp_{des}$  which represents the individually, quantified desired expectations. The resulting model is displayed in Fig. 13b. This extended model has two input parameters: the technical quality via downlink bandwidth in kBit/s and the individual desired expectation. Obviously, a lower desired expectation (5) results in a higher MOS score compared to higher expectations (1) (assuming that the technical quality is equal). The second line of Table 7 provides some additional information: the adjusted  $R^2$  value is higher compared to the pure QoS/QoE model and the RMSE is lower. To get the information about the added explanatory value of the factor  $exp_{des}$ , the squared Pearson correlation coefficient is calculated between the residuals of the pure QoS/QoE model and the factor  $exp_{des}$ . Hence, by including the information about desired expectations the MOS prediction accuracy was enhanced by 4.44%.

### Modelling of adequate expectations

In Study ASMO 3, our test participants had to download a 50 MB file via a website (see Fig. 14b) and with three applied different downlink bandwidth values: 4, 14 and 45 MBit/s. After the tasks were completed, the users filled out a questionnaire to get their quality impression regarding the perceived speed of the internet connection (“How

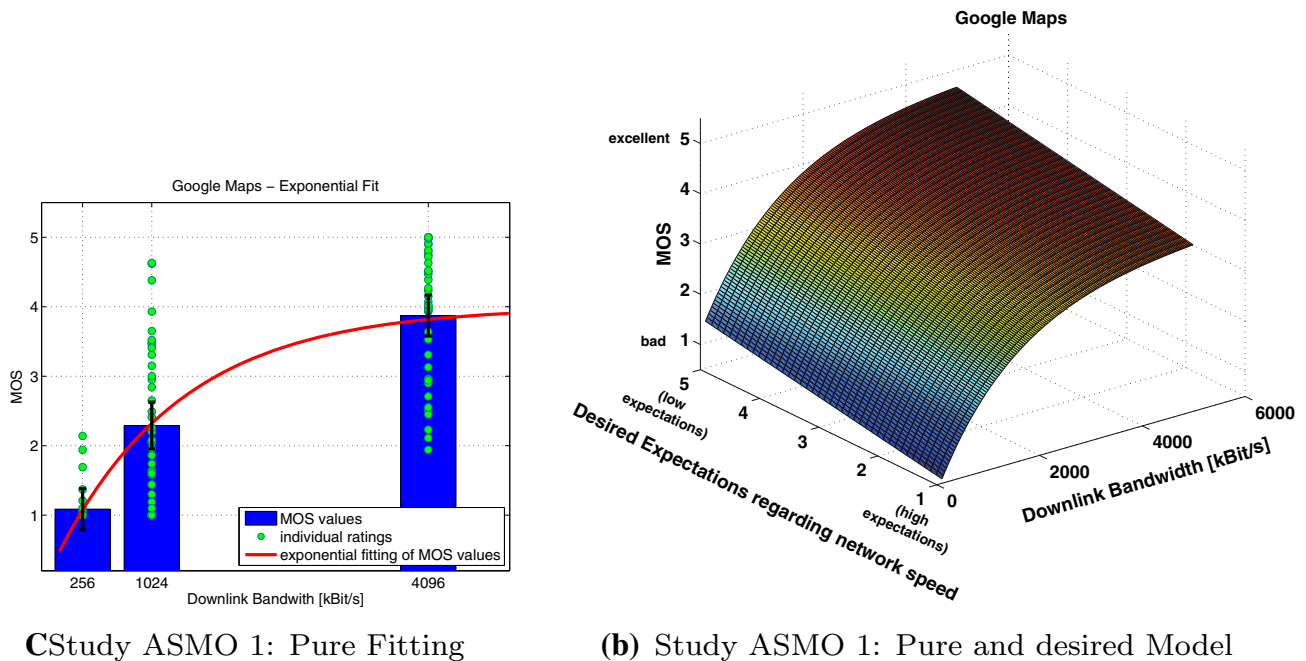


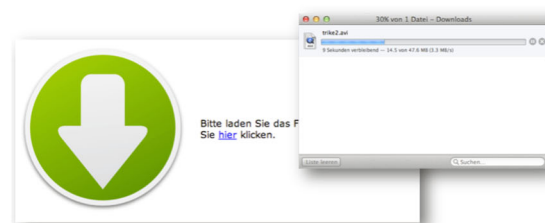
Fig. 13 Modelling of study ASMO 1: browsing Google maps

Table 7 Google maps models

Model type	Model	$R^2$	RMSE	Add. expl. value
QoS	$MOS_{Maps} = 4 - 3.52 * e^{-0.00072 * DLBW}$	0.6483	0.8386	–
QoS and $Exp_{des}$	$MOS_{Maps} = 3.73 - 3.59 * e^{-0.0007 * DLBW} + 0.21142 * exp_{des}$	0.6591	0.8256	+4.44 %



(a) Study ASMO 1: User Google Maps Task



(b) Study ASMO 3: User Download Task

Fig. 14 User tasks for study ASMO 1 and ASMO 3 (color figure online)

do you perceive the speed of the Internet connection?” with answer options ranging from 1 = bad to 5 = excellent). Additionally, we gathered information about adequate expectations with a questionnaire which includes the question “How long should it take to download a 50 MB file at home?” with the answer options 10, 30, 60 and 90 s, see “Assessment of Adequate Expectations” for details.

Figure 15a depicts the individual quality ratings (green dots), the resulting MOS values (blue bars) and the resulting exponential fitting curve. The first line of Table 8

shows the pure QoS/QoE model with the resulting adjusted  $R^2$  value (0 = no fitting of the model with the underlying data; 1 = perfect fitting) and the root-mean-square deviation.

Next, the pure QoS/QoE-model was extended with the additional additive, linear factor  $exp_{ade}$  which represents the individually, quantified adequate expectations. The resulting model is displayed in Fig. 15b. This extended model has two input parameters: the technical quality via downlink bandwidth in MBit/s and the individual

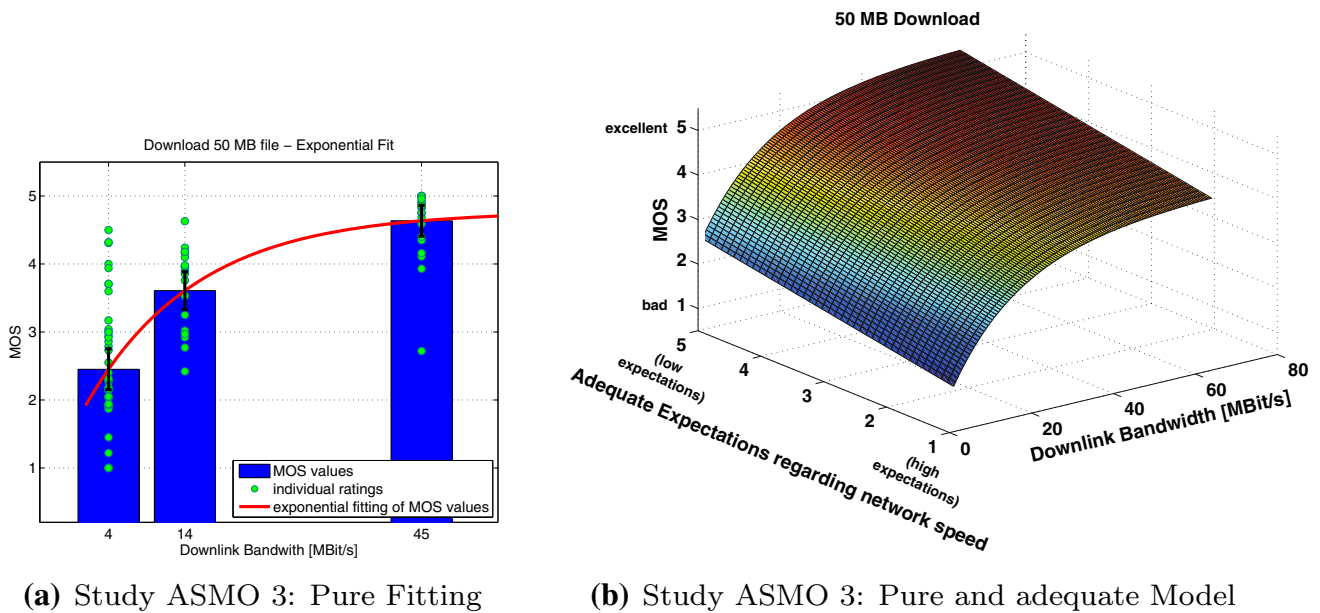


Fig. 15 Modelling of study ASMO 3: download tasks

Table 8 Study ASMO 3–50 MB download models

Model type	Model	$R^2$	RMSE	Add. expl. value
QoS	$MOS_{DL} = 4.77 - 3.06 * e^{-0.07 * DLBW}$	0.57	0.80	–
QoS and $Exp_{ade}$	$MOS_{DL} = 4.17 - 3.09 * e^{-0.06 * DLBW} + 0.26 * exp_{ade}$	0.61	0.77	+9.41%

adequate expectation. Obviously, a lower adequate expectation (5) results in a higher MOS score compared to higher adequate expectations (1). The second line of Table 8 provides some additional modeling information: the adjusted  $R^2$  value is higher compared to the pure QoS/QoE model and the RMSE is lower. To get the information about the added explanatory value of the factor  $exp_{ade}$ , the squared Pearson correlation coefficient is calculated between the residuals of the pure QoS/QoE model and the factor  $exp_{ade}$ . Hence, by including the information about adequate expectations the MOS prediction accuracy was enhanced by 9.41%.

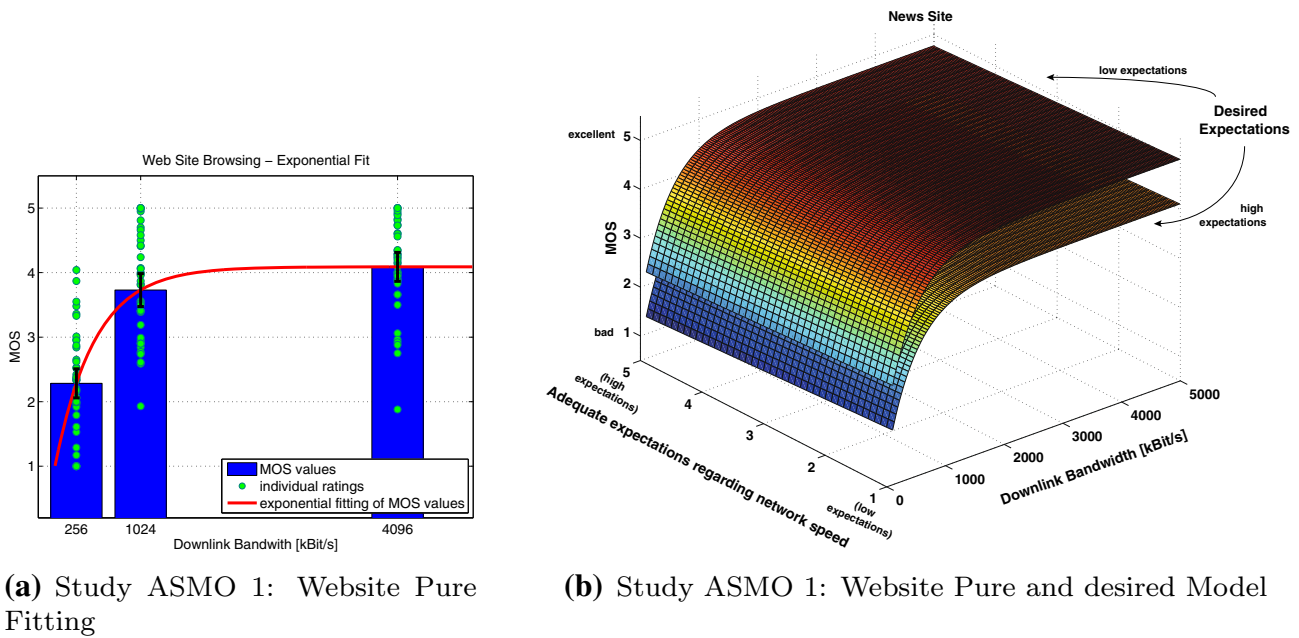
#### Modelling of adequate and desired expectations

In Study ASMO 1, among others our test participants had to browse a News Site (<http://www.nachrichten.yahoo.de>) via three different downlink bandwidth values: 256, 1024, 4096 kBit/s. After the tasks were completed the users filled out a questionnaire to get their quality impression regarding the perceived speed of the internet connection (“How do you perceive the speed of the Internet connection?” with answer options ranging from 1 = bad to 5 = excellent). Additionally, we gathered information about desired

expectations with ranking questions (see “Assessment of Desired Expectations” for more details). Finally, we use the position of the ranked element “Importance of Network Speed”, which value was between 1 (generally very important) and 5 (generally not important) to determine the desired expectation. Furthermore, we gathered information about adequate expectations with a questionnaire which includes the question “How fast should your Internet connection be at home for browsing the web, e.g., a News Site” with the answer options 0.256 Mbit/s, 0.512 Mbit/s, 1 Mbit/s, 4 Mbit/s and 8 Mbit/s. Note: In contrast to the previous “Assessment of Adequate Expectations”, the adequate expectations are related to the expected downlink bandwidth and not to a certain duration e.g. how long it should take to download/render a Web site.

Figure 16a depicts the individual quality ratings (green dots), the resulting MOS values (blue bars) and the resulting exponential fitting curve. The first line of Table 9 shows the pure QoS/QoE model with the resulting adjusted  $R^2$  value (0 = no fitting of the model with the underlying data; 1 = perfect fitting) and the root-mean-square deviation.

Next, the pure QoS/QoE-model was extended with the additional additive, linear factor  $exp_{des}$  which represents the individually, quantified desired expectations and the



**Fig. 16** Modelling of study ASMO 1: browsing a news site (color figure online)

**Table 9** Google maps models

Model type	Model	$R^2$	RMSE	Add. expl. value
QoS	$MOS_{News} = 4.09 - 3.09 * e^{-0.91016 * DLBW}$	0.497	0.779	–
QoS and $Exp_{des}$ and $Exp_{ade}$	$MOS_{News} = 4.06 - 3.09 * e^{-0.0021 * DLBW} - 0.066 * exp_{ade} + 0.228 * exp_{des}$	0.553	0.734	+12.42%

additional additive, linear factor  $exp_{ade}$  which represents the individually, quantified adequate expectations. The resulting model is displayed in Fig. 16b. This extended model has three input parameters: the technical quality, i.e., downlink bandwidth in kBit/s, the individual adequate and the individual desired expectation. Regarding legibility, in Fig. 16b, the resulting model for only two desired expectation values—lowest and highest possible values—is displayed. In contrast to the previous models, the overall added explanatory value +12.42% is the sum of two values: the squared Pearson correlation coefficient between the residuals of the pure QoS/QoE model and the factor  $exp_{des}$  and the squared Pearson correlation coefficient between the residuals of the pure QoS/QoE model and the factor  $exp_{ade}$ .

*Conclusions about expectation modelling*

In this section we have shown that it is expedient to include information about individual quality expectation in quantitative QoE models to enhance accuracy of MOS prediction. Depending on the expectation information—desire, adequate or both—the prediction enhancement is between 4.44 and 12.42%.

**Conclusions and future work**

In this article we have shown how expectations can systematically be integrated into QoE-related research. Based on our literature study regarding expectations in the fields of psychology, service quality and consumer satisfaction theory, we have extended an existing, fine-grained conceptual QoE model by the inclusion of *desired* and *adequate expectations* in the quality perception process. Further, dedicated questionnaires were proposed for direct expectation assessment. Finally, this information about individual expectations was used to improve quality prediction models, as shown in this paper.

Hence, we were able to prove that: (1) It is important to control and to describe experimental settings to avoid unintentional expectation triggering. (2) If expectations are to be included explicitly in an experiment, the triggering mechanism needs to be convincing and pellucid. (3) When using the proposed two short questionnaires in quality-related experiments, individual desired and adequate expectations can be quantified and the results can be utilized to improve QoE models with little effort. So, QoS/QoE management systems may benefit from this kind of enhanced models. Nevertheless, a more practical approach

is required to get expectation-related information from a large amount of end users. For examples, records of previous usage behaviour could be used to derive information of QoS/QoE-relevant expectations. Field trials might be an appropriate method to conduct such kind of research.

Nevertheless, several open questions and challenges emerged during our research, and therefore, further investigation is necessary to broaden our knowledge about the complex interplay between expectations and quality:

- From a *theoretical* point of view, the presented results and methods are primarily based on pragmatic approaches which provide practical tools and methods for including expectations in perceptual (“subjective”) and instrumental (“objective”) QoE assessment. Future work should complement our research with an explicit, deep analysis of the psychology behind QoE-related expectations. Hence, it might be necessary to develop and evaluate models focusing on psychological aspects to get a better understanding of how expectations are mentally set, adapted over time and how expectations interact with quality perception in a detailed way.
- *Methodologically*, our questionnaires regarding desired and adequate expectations should be adapted and improved to get more precise and valid information. For example, our questionnaires about adequate expectations were designed for certain situations, e.g., the evaluation of expected downlink bandwidth in a mobile Internet usage scenario. This approach requires some kind of abstraction by the involved users and for some use cases this approach might not be appropriate. For example, it is undefined which is the most appropriate quality feature regarding video quality ought to be included in the adequate expectations questionnaire. In the studies presented in this journal paper the participants had to state expectations regarding video bit rate and video resolution. Nevertheless, less abstract features are required, e.g., regarding Web QoE it is reasonable to ask about indirect quality indicators like page load times or file download times. One promising approach could be to provide examples of different video qualities and the user has to select an adequate quality. Additionally, it might be interesting to extend the adequate expectation questionnaire with questions regarding the specific (e.g. downlink bandwidth, video resolution, etc.) barely accepted quality as well as the actually desired quality-level instead of just inquiring a single value, which might be useful for advancing quantitative QoE models.
- Furthermore, also on a methodological level, we have to emphasize that in our studies we focused on the assessment of expectations *prior* to a test condition or task. In this respect, future work should investigate how (the impact of) expectations can be reliably assessed after the fact, e.g. by inquiring subjects on actual fulfillment of expectations. However, the challenge here is such questions directly interfere with the quality ratings that QoE test subjects are expected to provide at the same stage, i.e. MOS ratings are likely to strongly correlate with degree-of-expectations-fulfillment ratings. Thus, dedicated studies would be required which distribute these two types of questions across different control groups.
- In addition, we have to note that our empirical research so far has been deliberately based on lab-based studies in order to obtain results in a controlled environment. A logical next step would be to complement this work by means of field or crowdsourcing studies conducted in real-world settings. In principle, the same methods presented in this article can be applied here (ranking questionnaires, etc.), particularly in order to deeper investigate differences stated expectations and actual user behaviour. In the light of the previous point, we recommend to explicitly ask participants about their expectations not too often (e.g. only at beginning and end of the trial) in order to avoid biasing perception and judgement. Furthermore, we have to distinguish between small-scale and large-scale field trials. Small-scale field trial represents the natural extension of lab studies in the way that a relatively small sample is recruited, which would still allow for the necessary deep investigation on participants’ expectation profiles (influenced by personality, motives, etc.) via interviews (and similar methods) and relate them to their behaviour patterns, quality ratings, etc., however based on real-world situations and contexts. Armed with the resulting models and relationships, large-scale field trials (resembling the surveys and campaigns typically conducted by carriers) can be then conducted in order to practically apply these models not only to improved QoE prediction but also to prediction of (economic) behaviour like service usage intensity and churn. However, due to the significantly high number of users involved (e.g. all subscribers of an operator), these large-scale field trials required other approaches of collecting information about user expectations must be evaluated. For example, information about expectations could be derived to some extent from existing socio-economical data (sex, age, usage habits, etc.), a step which only can be performed, if a validated model from a small-scale field trial has been developed.
- As regards *modelling* we extended fitting-based models with linear factors representing expectation-information to demonstrate their usefulness. Here, other, non-linear integration approaches could lead to even more promising results. Moreover, the modelling approaches

that may benefit from explicit inclusion of expectations are not restricted to such closed-form formula-based approaches, but include modelling using machine learning (see “[Extending quantitative QoE models with Information about Expectations\(RQ3\)](#)”).

- So far, several *additional expectation-related factors* were neglected in our work. As one important aspect, the link between individual quality features or dimensions and the *context* in which these occur or the technology that is being applied and thus the underlying expectations needs to be investigated. For example, in case of speech, background noise from a distant speaker (related to the dimension “noisiness”) might be more easily accepted when it is known that this speaker is located in a public place. We also neglected the impact of specific *tasks* on quality expectations. For example, if music is consumed in the background e.g. during cooking activities, quality expectations are likely to be different compared to a situation in which music is intensively consumed as a foreground task via a Hi-Fi setup. In QoE research, the influence of tasks on quality perception has been examined e.g. Strohmeier et al. (2012), but it is still unclear how specific tasks impact individual expectations. Furthermore, the results of study 1 and 2 show that expectations *adapt over time*, cf. Fig. 9. So far, only little research has been carried out to investigate which factors influence modifications of expectations over time. One could assume that adapted usage behavior (e.g. changing from a 3G to a LTE Internet connection on the smartphone) and media consumption (e.g. advertisements, which has been demonstrated by Higgs et al. 2005) shape the expectations, but empirical evidence directly collected in the domain of QoE research is currently missing. We used *DiffMOS* and Confidence Intervals combined with CDF-plots to verify if our expectation triggering has a significant effect. In this context, our CDF plots—e.g. see Fig. 9a—show that in most cases a positive *and* a negative effect of a certain trigger occur. Hence, more research is needed to clarify if individual factors (for example gender, age, previous usage behavior etc.) could be used to explain these heterogenous responses.

In general, we encourage the interested reader to take up our approaches in order to evaluate, adapt and extend our methods in order to include user expectations in QoE research. For example, for certain scenarios our questionnaires need to be modified, e.g., if adequate expectations are collected for video quality, how should participants be asked regarding perceived quality features? In case of Web QoE, it is possible to ask about quality features such as perceived page load times or download times. For

multimedia applications such as telephony, conferencing or video streaming, an analysis in terms of quality dimensions could be carried out in parallel tests, using the typically applied techniques such as Multi-dimensional Scaling (MDS, see Carroll 1972), Attribute Scaling/Semantic Differential (see Osgood et al. 1957) or Open Profiling of Quality, Strohmeier et al. (2010). In subsequent preference mapping, different mapping-parameter settings are conceivable that reflect the specific expectations in a given context. Here too, direct questionnaires may be conceivable, asking for aspects such as resolution, certain observed or heard errors, etc.

In this sense, this article aims to serve as foundation and trigger to encourage novel research on expectations, not only in order to better assess and quantify them, but also to gain a better understanding of human quality formation processes and thereby ultimately enable creation of better ICT products and services.

## Appendix: Questionnaires

The paper questionnaires were originally generated in German. Regarding further use and in the context of this thesis, translated versions are included in the appendix. Only the relevant parts of the questionnaires are depicted, i.e., test numbers, pagination, test user ID, etc. are omitted (Figs. 17, 18, 19, 20, 21).

To what extent do you agree or disagree with the following statements:

	full applies	rather applies	neither nor	rather not applies	not applies
Mobile Internet access is standard.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I expect 100% stability from my mobile Internet access.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mobile high-speed Internet access is very important	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fixed Internet access is standard.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
I expect 100% stability from my fixed Internet access.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Fixed high-speed Internet access is very important	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Fig. 17** Study TRIG 1 and TRIG 2—Questionnaire about the assessment of general expectations w.r.t. fixed and mobile internet access (color figure online)

**Fig. 18** Study ASMO 1—  
Questionnaire about desired  
expectations w.r.t. Internet  
access features

What do you expect from your Internet connection? Please order the following features of an Internet connection regarding your personal importance. (1=most important, 2=second important, etc.)

- High Speed
- Low monthly fee
- Short Contract duration
- Good Support via E-Mail and telephone
- Unlimited download volume

**Fig. 19** Study ASMO 2—  
Questionnaire about desired  
expectations w.r.t. Video on  
Demand providers

What do you expect from a Video on Demand Provider? Explanation: Via a Video on Demand Service it is possible to consume movies and TV Series over the Internet on your TV for a certain fee. The videos are rent and streamed, i.e., the videos are not bought.

What do you expect from a Video on Demand provider? Please order the following regarding your personal importance. (1=most important, 2=second important, etc.)

- Large amount of available movies and TV series
- Good support via E-Mail and telephone
- Low monthly fee
- High video quality
- Short contract duration

**Fig. 20** Study ASMO 3—  
Questionnaire about adequate  
expectations w.r.t. mobile/fixed  
Internet connections

How fast should your **mobile Internet connection** be (data stick connected to your Laptop) while you browse a typical **Internet page** like maps.google.at or youtube.com?

- 1 MBit/s    4 MBit/s    14 MBit/s    30 MBit/s    45 MBit/s

The reply to this question was:

Very easy              very difficult (=I could not answer this question)

How long should it maximally take to **download a 50 MB** file via your **mobile Internet connection** (data stick connected to your Laptop)?

- 10 seconds    30 seconds    1 minute    1 minute 30 seconds    2 minutes

The reply to this question was:

Very easy              very difficult (=I could not answer this question)

**How long** should it take **maximally** to **download an Internet page completely** (at home via ADSL/Cable,etc.)?

- 2 seconds    3-7 seconds    8-12 seconds    13-17 seconds    more than 17 Seconds

The reply to this question was:

Very easy              very difficult (=I could not answer this question)

How fast should your **Internet connection** be at **home** (ADSL, Cable, etc.) while you **download a 50 MB** file?

- 1 MBit/s    4 MBit/s    14 MBit/s    30 MBit/s    45 MBit/s

The reply to this question was:

Very easy              very difficult (=I could not answer this question)



**Fig. 21** Study ASMO 3—  
Questionnaire about adequate expectations  
w.r.t. mobile/fixed Internet connections

How fast should your **Internet connection** at home (ADSL, Cable, etc.) be while you browse a typical **Internet page** like maps.google.at or youtube.com?

1 MBit/s  4 MBit/s  14 MBit/s  30 MBit/s  45 MBit/s

The reply to this question was:

Very easy      very difficult (=I could not answer this question)

How long should it maximally take to **download a 50 MB** file via your **Internet connection at home** (ADSL, Cable, etc.)?

10 seconds  30 seconds  1 minute  1 minute 30 seconds  2 minutes

The reply to this question was:

Very easy      very difficult (=I could not answer this question)

**How long** should it take **maximally to download an Internet page completely** with your mobile Internet connection (data stick connected to your Laptop)?

2 seconds  3-7 seconds  8-12 seconds  13-17 seconds  more than 18 seconds

The reply to this question was:

Very easy      very difficult (=I could not answer this question)

How fast should your **mobile Internet connection** (data stick connected to your Laptop) be while you **download a 50 MB** file?

1 MBit/s  4 MBit/s  14 MBit/s  30 MBit/s  45 MBit/s

The reply to this question was:

Very easy      very difficult (=I could not answer this question)

## References

- Aguiar E, Riker A, Abelém A, Cerqueira E, Mu M (2012) Video quality estimator for wireless mesh networks. In: Proceedings of the 2012 IEEE 20th international workshop on quality of service, IWQoS '12. IEEE Press, Piscataway, pp 1:1–1:9. <http://dl.acm.org/citation.cfm?id=2330748.2330749>
- Albert W, Dixon E (2003) Is this what you expected? The use of expectation measures in usability testing. In: Proceedings of usability professionals association conference, Scottsdale, AZ
- Bacioccola A, Cicconetti C, Stea G (2007) User-level performance evaluation of VoIP using ns-2. In: ValueTools '07 proceedings of the 2nd international conference on Performance evaluation methodologies and tools, Brussels, Belgium
- Bonito JA, Burgoon JK, Bengtsson B (1999) The role of expectations in human–computer interaction. In: Proceedings of the international ACM SIGGROUP conference on supporting group work, GROUP '99. ACM, New York, pp 229–238. doi:[10.1145/320297.320324](https://doi.org/10.1145/320297.320324)
- Boulding W, Kalra A, Staelin R, Zeithaml VA (1993) A dynamic process model of service quality: from expectations to behavioral intentions. *J Mark Res* 30(1):7–27
- Buttle F (1996) Servqual: review, critique, research agenda. *Eur J Mark* 30(1):8–32
- Cadotte ER, Woodruff RB, Jenkins RL (1987) Expectations and norms in models of consumer satisfaction. *J Mark Res* 24(3):305–314. <http://www.jstor.org/stable/3151641>
- Carroll JD (1972) Individual preferences and multidimensional scaling. In: Shepard RN, Romney AK, Nerlove SB (eds) *Multidimensional scaling: theory and applications in the behavioral sciences*, vol 1, pp 105–155
- Cerqueira E, Veloso L, Curado M, Monteiro E, Mendes PM (2008) Quality level control for multi-user sessions in future generation networks. In: IEEE GLOBECOM global telecommunications conference, New Orleans, LO, USA
- Egger S, Reichl P, Hofffeld T, Schatz R (2012) Time is bandwidth? Narrowing the gap between subjective time perception and quality of experience. In: 2012 IEEE international conference on communications (ICC), pp 1325–1330. doi:[10.1109/ICC.2012.6363769](https://doi.org/10.1109/ICC.2012.6363769)
- Gerrig RJ (2012) *Psychology and life*, 20th edn. Pearson, New Jersey
- Hartikainen M, Salonen E-P, Turunen M (2004) Subjective evaluation of spoken dialogue systems using servqual method. In: *Inter-speech 2004—ICSLP*, 8th international conference on spoken language processing, Jeju Island, Korea
- Higgs B, Polonsky MJ, Hollick M (2005) Measuring expectations: forecast vs. ideal expectations. Does it really matter? *J Retail Consum Serv* 12(1):49–64
- Hirth M, Hofffeld T, Tran-Gia P (2011) Anatomy of a crowdsourcing platform—using the example of microworkers.com. In: *Workshop on future internet and next generation networks (FINGNet)*, Seoul, Korea
- Hofffeld T, Keimel C (2014) *Crowdsourcing in QoE evaluation. In: Quality of experience. Advanced concepts, applications and methods*. Springer, Berlin

- Hohlfeld O, Pujol E, Ciucu F, Feldmann A, Barford P (2014) A QoE perspective on sizing network buffers. In: Proceedings of the 2014 conference on internet measurement conference, Vancouver, BC, Canada
- Hsieh Y-H, Yuan S-T (2009) A quantitative approach to measure customer expectation for service innovation within service experience delivery. In: Proceedings of the 11th international conference on electronic commerce, ICEC '09. ACM, New York, pp 38–46. doi:[10.1145/1593254.1593260](https://doi.org/10.1145/1593254.1593260)
- Huong P-T, Hiep H-V, Miyoshi T, Yamazaki T (2014) QoE-driven bandwidth allocation method based on user characteristics. In: Proceedings of the 16th network operations and management symposium (APNOMS), Taiwan
- International Telecommunication Union (2005) The e-model, a computational model for use in transmission planning. Technical report, ITU-T Recommendation G.107
- International Telecommunication Union (2014) Estimating end-to-end performance in ip networks for data applications, ITU-T Recommendation G.1030
- Jekosch U (2006) Voice and speech quality perception: assessment and evaluation. Springer, Berlin
- Kim DJ, Ferrin DL, Rao HR (2003) A study of the effect of consumer trust on consumer expectations and satisfaction: the Korean experience. In: Proceedings of the 5th international conference on electronic commerce, ICEC '03. ACM, New York, pp 310–315. doi:[10.1145/948005.948046](https://doi.org/10.1145/948005.948046)
- Kristensen K, Martensen A, Gronholdt L (1999) Measuring the impact of buying behaviour on customer satisfaction. *Total Qual Manag* 10(4–5):602–614. doi:[10.1080/0954412997587](https://doi.org/10.1080/0954412997587). <http://www.tandfonline.com/doi/abs/10.1080/0954412997587>
- Le Callet P, Möller S, Perks A (2012) Qualinet white paper on definitions of quality of experience. Technical report, European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003) (2012, Lausanne, Switzerland)
- McGill AL, Iacobucci D (1992) The role of post-experience comparison standards in the evaluation of unfamiliar services. *Adv Consum Res* 19:570–578
- Menkovski V, Oredope A, Liotta A, Sánchez AC (2009) Predicting quality of experience in multimedia streaming. In: Proceedings of the 7th international conference on advances in mobile computing and multimedia, MoMM '09. ACM, New York, pp 52–59. doi:[10.1145/1821748.1821766](https://doi.org/10.1145/1821748.1821766). <http://doi.acm.org/10.1145/1821748.1821766>
- Menkovski V, Exarchakos G, Liotta A (2010) Online QoE prediction. In: Second international workshop on quality of multimedia experience (QoMEX), Trondheim, pp 118–123. doi:[10.1109/QOMEX.2010.5517692](https://doi.org/10.1109/QOMEX.2010.5517692)
- Micheli D, Curto A, Mazzurco M, Pierucci L, Fantacci R, Simoni R (2013) Handset and network quality performance benchmarking for QoE improvement. In: Globecom workshops (GC Wkshps), Atlanta, GA, USA
- Mitra K, Zaslavsky A, Åhlund C (2011) A probabilistic context-aware approach for quality of experience measurement in pervasive systems. In: Proceedings of the 2011 ACM symposium on applied computing, TaiChung, Taiwan
- Möller S (2000) Assessment and prediction of speech quality in telecommunications. Springer US, New York
- Möller S, Raake A (eds) (2014) Quality of experience: advanced concepts, applications and methods. Springer, Cham. doi:[10.1007/978-3-319-02681-7](https://doi.org/10.1007/978-3-319-02681-7)
- Monath T, Kind M, Heger T, Schlesinger M, Aznar J (2010) Economical analysis of experience-optimized service delivery. In: 2010 9th Conference on Telecommunications Internet and Media Techno Economics (CTTE), Ghent
- Munson MJ, McIntyre SH (1979) Developing practical procedures for the measurement of personal values in cross-cultural marketing. *J Mark Res* 16(1):48–52
- Nicolas S, Van den Broeck W, Pitrey Y, Vermeulen B, Demeester P (2012) Lessons learned during real-life QoE assessment. In: 10th European conference on interactive TV and video (Euro ITV-2012), Ghent University, Department of Information Technology
- Ojasalo J (2001) Managing customer expectations in professional services. *Manag Serv Qual* 11(3):200
- Oliver RL (1996) Satisfaction: a behavioral perspective on the consumer. Irwin/Mc Graw-Hill, New York
- O'Neil M, Palmer A (2003) An exploratory study of the effects of experience on consumer perceptions of the service quality construct. *Manag Serv Qual* 13(3):187–196
- Osgood C, Suci G, Tannenbaum P (1957) The measurement of meaning. University of Illinois Press, Urbana
- Parasuraman A, Zeithaml VA, Berry LL (1988) SERVQUAL: a multiple-item scale for measuring consumer perceptions of service quality. *J Retail* 64(1):12–40. <http://search.ebscohost.com/login.aspx?direct=true&#38;db=buh&#38;AN=6353260&#38;site=ehost-live>
- Pécharde S, Carnec M, Le Callet P, Barba D (2006) From SD to HD television: effects of h. 264 distortions versus display size on quality of experience. In: IEEE international conference on image processing. IEEE, Atlanta, pp 409–412
- Plücker R (1998) Untersuchung des Einflusses der Qualitätserwartung und der Telefonier-Situation auf die Beurteilung von Telefonsprache. Ruhr Universität Bochum, Study project report (unpublished), Institut für Kommunikationsakustik
- Raake A (2006) Speech quality of VoIP—assessment and prediction. Wiley, Chichester
- Raake A, Egger S (2014) Quality and quality of experience. In: Quality of experience. Advanced concepts, applications and methods. Springer, Berlin
- Reeves CA, Bednar DA (1994) Defining quality: alternatives and implications. *Acad Manag Rev* 19(3):419–445
- Reiter U, Brunnström K, De Moor K, Larabi M-C, Pereira M, Pinheiro A, You J, Zgank A (2014) Factors influencing quality of experience. In: Quality of experience. Advanced concepts, applications and methods. Springer, Berlin
- Riedel J (1998) Einfluß der Erwartungshaltung des Benutzers auf die Qualitätsbeurteilung von Telefonübertragungen an Internet-Terminals. Study project report (unpublished), Institut für Kommunikationsakustik, Ruhr-Universität, Deutschland Bochum
- Rokeach M (1968) Beliefs, attitudes, and values: a theory of organization and change. Jossey-Bass, San Francisco
- Rust RT, Oliver RL (1994) Service quality: insights and managerial implications. In: Service quality: new directions in theory and practice. Sage Publications, California
- Sackl A, Egger S, Schatz R (2013) Where's the music? Comparing the QoE impact of temporal impairments between music and video streaming. In: Fifth international workshop on quality of multimedia experience (QoMEX), pp 64–69. doi:[10.1109/QoMEX.2013.6603212](https://doi.org/10.1109/QoMEX.2013.6603212)
- Schatz R, Hoßfeld T, Janowski L, Egger S (2013) From packets to people: Quality of experience as a new measurement challenge. In: Biersack E, Callegari C, Matijasevic M (eds) Data traffic monitoring and analysis, vol 7754. Lecture notes in computer science. Springer, Berlin, pp 219–263
- Shirai M, Meyer R (1997) Learning and the cognitive algebra of price expectations. *J Consum Psychol* 4:365–388
- Sternberg RJ (2011) Cognitive psychology, 6th edn. Wadsworth Publishing, Belmont
- Strohmeier D, Jumisko-Pyykkö S, Kunze K (2010) Open profiling of quality: a mixed method approach to understanding multimodal quality perception. In: Advances in multimedia, vol 2010, p 28

- Strohmeier D, Jumisko-Pyykko S, Raake A (2012) Toward task-dependent evaluation of web-QoE: free exploration vs. who ate what? In: Globecom workshops (GC Wkshps), 2012 IEEE, pp 1309–1313. doi:[10.1109/GLOCOMW.2012.6477771](https://doi.org/10.1109/GLOCOMW.2012.6477771)
- Szajna B, Scamell RW (1993) The effects of information system user expectations on their performance and perceptions. *MIS Q* 17(4):493–516. <http://www.jstor.org/stable/249589>
- Vandenbroucke K, Ferreira D, Goncalves J, Kostakos V, De Moor K (2014) Mobile cloud storage: a contextual experience. In: Proceedings of the 16th international conference on human-computer interaction with mobile devices and services, Toronto, ON, Canada
- Wältermann M, Raake A, Möller S (2010) Quality dimensions of narrowband and wideband speech transmission. *Acta Acust United Acust* 96(6):1090–1103
- Watson RT, Pitt LF, Kavan CB (1998) Measuring information systems service quality: lessons from two longitudinal case studies. *MIS Q* 22(1):61–79
- Wijnants M, Agten S, Quax P, Lamotte W (2009) Investigating the relationship between QoS and QoE in a mixed desktop/handheld gaming setting. In: Co-next student workshop '09: proceedings of the 5th international student workshop on emerging networking experiments and technologies, Rome, Italy
- Wilkie WL (1994) *Consumer behaviour*, 3rd edn. Wiley, Canada, p 118
- Yajun H, Wenan Z, Yu D (2014) Research on the user behavior-based QoE evaluation method for http mobile streaming. In: Proceedings of the 10th international conference on broadband and wireless computing, communication and applications, Krakow, Poland
- Zeithaml VA, Parasuraman A, Berry LL (1990) *Delivering service quality: balancing consumer perceptions and expectations*. The Free Press, New York
- Zeithaml VA, Berry LL, Parasuraman A (1993) The nature and determinants of customer expectations of service. *J Acad Mark Sci* 21(1):1–12. doi:[10.1177/0092070393211001](https://doi.org/10.1177/0092070393211001)
- Zielinski S, Rumsey F, Bech S (2008) On some biases encountered in modern audio quality listening tests—a review. *J Audio Eng Soc* 56(6):427–451. <http://www.aes.org/e-lib/browse.cfm?elib=14393>