Evaluating the impact of expectations on end-user quality perception

Andreas Sackl, Raimund Schatz

Telecommunications Research Center Vienna (FTW), Vienna, Austria
Email: {sackl | schatz}@ftw.at

Abstract

In Quality of Experience (QoE) research, the concept of 'expectations' is frequently used but seldom defined with sufficient clarity and precision. In this paper, we present theoretical and conceptual approaches originating from various research fields that deal with expectations and its measurement. In addition, we also present results from our own empirical work to exemplify how to measure and to evoke expectations in the context of perceived quality for multimedia services. Together these contributions form the basis for an interdisciplinary discussion of the nature of expectations on quality perception as well as its role in QoE research.

Index Terms: Quality of Experience, Expectations, Assessment Methodologies

1. Introduction

The term *expectations* is frequently used in the context of quality perception and Quality of Experience (QoE) related research. Its relevance is also confirmed by its appearance in the context of various definitions and models of quality: Firstly, in the Qualinet QoE definition white paper expectations are described as key factor determining the end user's perceptions and resulting 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 users personality and current state."[1]*

Secondly, even ITU-T Recommendation P.10 highlights the relevance of user expectations, despite its primarily technical focus. 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."[2]*

Thirdly, QoE-models like the one used in [3] explicitly include expectations as a main user influence factor (see Figure 1).

Nevertheless, most of the existing practical and theoretical QoE research literature does not provide an exact definition or operationalization of expectations, nor does it address questions of how to measure the influence of expectations-related of user aspects on quality
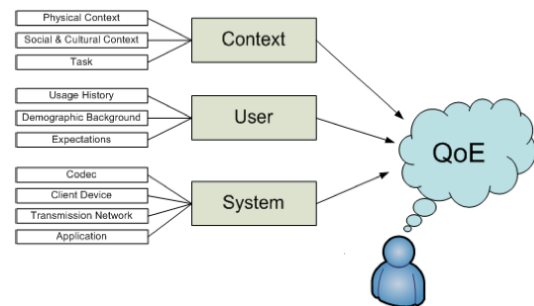

Figure 1: The term expectations as an important user-related influence factor on QoE.

perception. However, given the importance of expectations in the context of QoE definitions and frameworks, it is paramount to have a clear common understanding of the concept and how it can be addressed in experimental research. But before dealing with the *influence* of expectations on QoE, the *emergence*, the *measurability* and the *deliberate triggering* of specific expectations have to be understood and examined.

The rest of the paper is structured as follows: In the second section we discuss some theoretical work from various disciplines, e.g. (socio-) psychology and Human-Computer-Interaction, and their approaches to define and include expectations in the respective field. In the subsequent section results from our empirical user studies dealing with the influence of expectations on QoE are presented. In section 4, we present our conclusions.

2. Related Work

Outside the domain of QoE research, the concept of expectations has been examined in various fields to a fairly broad extent, e.g. HCI, economics, psychology, etc. The authors of [4] define expectations as *"a kind of schemata that focuses interpretation processes on specific meanings and functions of communicative action"[4]*. They also state that expectations provide a basis for evolution of human behavior which is crucial in situations which might lead to a violation of expectations. In expectancy violation theory (EVT)[5], three sources are described which form expectations in an interaction sce-

nario: 1) socio-demographic aspects, language style of the opponent and appearance, 2) The nature of the relationship (similarity, attraction, power differentials etc.) and 3) the context of the relationship like the nature of the task influences the generated expectation. According to EVT, expectations can be violated in a negative *and* positive way, i.e. there is not only a non-fulfillment but also a (positively experienced) over-fulfillment.

In the context of information system evaluation, [6] have used the SERVQUAL method [7] to measure existing expectations and to compare it with gained experiences: Test users had to indicate their *ideal* information system and to answer 22 questions, each had to be rated via a 7-point scale (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 certain information system implementation and they had 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, [8] used the Expectation Confirmation Theory (ECT) [9] to measure the user-satisfaction of web services, which develops from a (customer's) comparison of post-purchase evaluation of a product or service with pre-purchase expectations. Following this theory, users form specific expectations regarding a desired product. After a trial phase, the users form perceptions about its performance. After that, the participants determine if their expectations were confirmed regarding the perceived performance. Then, the users satisfaction level is resulting from the previous confirmation and the underlying expectations. Finally, a reuse or repurchase is considered or not.

Furthermore, [10] uses the Weber/Fechner-law (which is well known in the context of QoE) to generate a quantitative expectation measurement model, which mathematically describes the relationship between the desired and the adequate expectations of customers regarding service providers. They distinguish between *desired expectations* and *adequate expectations*.

According to [11], the desired expectations are rather stable and invariant, e.g. some users are always concerned about high quality or low prices, so their basic needs do not change. Adequate expectations are more flexible and they are influenced by the context. Between these two kind of expectations, there is the so called *zone of tolerance*, i.e. if the *perceived* expectation is in between the invariable *desired* expectation and the variable *adequate* expectation, the user/customer *accept* the perceived service, see Fig. 2. As depicted in Fig. 1, current QoE models describe the influence of context and
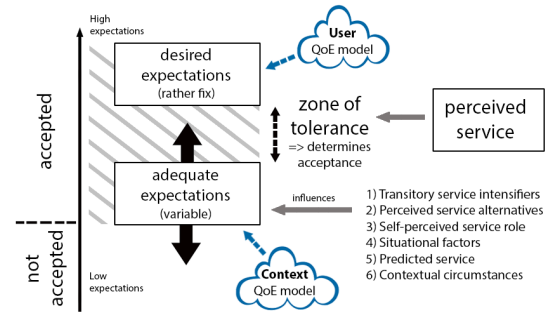

Figure 2: Desired and adequate expectations, based on [11]. If the perceived service is located within the zone of tolerance, the service quality is accepted

user variables on QoE. The term "user" perfectly matches with the nature of the relatively stable desired expectations, whereas the term "context" is compatible with the context-sensitive term adequate expectations. Therefore, the described expectations model in Fig. 2 is perfectly in line with established QoE models.

In the context of typical business/customer relations, adequate expectations are influenced by: (1) *transitory service intensifiers*, e.g. the urgency of a situation. (2) *perceived service alternatives*, i.e., if there are alternatives available in the current situation or it is possible to solve an issue without external support, the adequate expectations are getting 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, i.e., it is not always possible to blame others for non-fulfilled expectations. Therefore, the more pretentious the level of a customers view on the self-perceived service role is, the higher the level of adequate service, (4) *Situational factors* can lower the level of adequate service, e.g., environmental influences are independent from a service provider. Customers recognize, that this is not a fault of the provider and they accept a lower service level, (5) *Predicted service*, this is the service customers believe they are likely to get. (6) *Contextual circumstances* could be for example company controlled prices. A participant in the study of [11] stated, that a "price increases don't really drive up expectations. But my tolerance level will become more stringent/less flexible with an increase".

Summarizing, the nature of expectations has been examined by various research areas, e.g. HCI, economics, psychology, etc. and its below-mentioned findings are valuable to explore the role of expectations in the field of QoE:

- Expectations depend on a broad variance of influences, it is complex to understand how they emerge and how the influence quality perception

- Expectations can be negatively perceived as *under-fulfilled*, but also positivly perceived as *over-*

fulfilled. Therefore, QoE research should not to forget to test situations in which expectations are over-fulfilled (at the moment, the focus of research relates to situations in which expectations are not fulfilled)

- In other research fields, *questionnaires* are used to measure expectations. But this approach requires expectations which can be *verbalized* and which are *quantifiable*, which is not always possible in the field of QoE, e.g. it might be challenging for an average end user to verbalize her expectations on internet connection speed in Mbit/s.

- According to the perviously mentioned theoretical expectation theories, *prior experiences* are mandatory for under- or fulfillment of expectations.

- There is a relationship between relatively stable, higher *desired expectations* and variable, adequate expectations, which influences the acceptance of a certain service.

3. Empirical work

The complex nature of expectations requires separate approaches regarding examination. In this section, our empirical attempts are described to exemplify how expectations can be *triggered*, *measured* and how they *emerge*.

3.1. Case study: fixed vs. mobile internet access

To measure the influence of expectations in specific quality-related usage scenarios, it is necessary to handle expectations as an *independent variable* in the study design. Therefore, situations have to be generated which provoke certain user expectations, while all other (technical) settings remain constant.

We conducted two experiments: our first study (3G/ADSL study), which was executed in 2011, examined the influence of a wireline 3G and a fixed linw ADSL internet connection on quality perception. Our second study (LTE/ADSL study) was conducted in 2013 and used a similar setup like the previous study to examine the influence of the connection type on quality perception. In both experiments, our test users had to manually switch by themselves via a physical device the connection type several times during the test procedure [12]. In fact, both connection types (wireline/wireless) were identical, i.e. the 3G/LTE modem was only a non-functional, but realistic mock-up. For each technical condition (certain values for up/downlink bandwidth, delay, service), our users tried out both connection types and rated them via a standard MOS scale. By comparing these two MOS values we were able to measure how the connection type influences expectations and the quality perception.
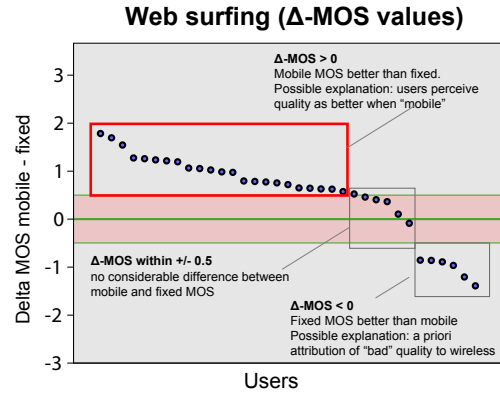

Figure 3: 3G/ADSL study: Δ MOS (Mean Opinion Score) values for each user regarding two access modes (3G vs. wireline) for browsing a website. The influence of expectations is visible for $|\Delta \text{MOS}| > 0.5$

3.1.1. Triggering expectations

Figure 3 shows that most of the users had lower expectations regarding using 3G in contrast to a wireline connection for browsing a news-site. In contrast to this, in the LTE/ADSL study we found out, that the difference between a wireline and a LTE connection is less pronounced, i.e., the MOS scores according to the set downlink bandwidth were more or less independent from the (technical identical) connection type. It seems that the users were less tolerant when accessing via LTE in contrast to 3G, i.e., the participants had the same, high expectations for LTE as for fixed line ADSL access.

3.1.2. Measurement of general expectations

We also explicitly asked our participant via a questionnaire about their expectations regarding internet access. Figure 4 shows a comparison between the 3G/ADSL and the LTE/ADSL study. Whereas the question regarding the necessity did not differ, the answers regarding reliability and high speed vary: In 2011, there was a clear difference regarding the expectations for high-speed access between the access technologies, i.e., for 3G high-speed was not mandatory. In 2013, the participants expect for both connection types a fast connection. These findings are in line with the MOS ratings, i.e., users made no difference between accessing a website via ADSL or 3G. Interestingly, the expectations regarding reliability for mobile access *dropped* over time. Whereas in 2011, the users imposed the same reliability standard for 3G as for ADSL, the users in 2013 were more tolerant regarding this quality aspect. One possible explanation is the varying private usage duration: in 2011, on average our test participants used mobile access for 15 month, whereas in 2013 the average duration of mobile access was 28
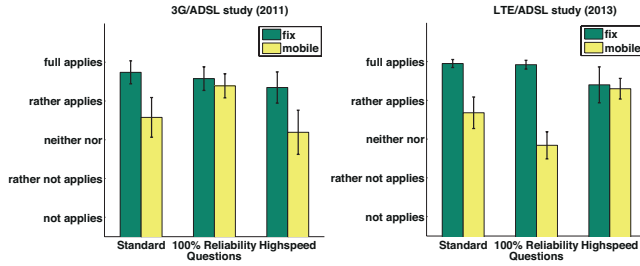
Figure 4: Measured general expectations. First question: "Mobile/fixed internet access is standard." Second question: "I expect 100% stability from my mobile/fixed internet access." Third Question: "Mobile/fixed high-speed internet access is very important"

month. So, the usage duration had almost doubled. Experienced low coverage or dead zones could shape existing expectations regarding mobile internet access, which leads to lower reliability expectations.

According to Figure 2 and [11], the desired expectations for internet access regarding speed and reliability changed over time. For mobile internet access speed, higher desired expectations lead to higher adequate expectations, i.e., the zone of tolerance is getting smaller[1].

3.1.3. Measurement of specific expectations

For the LTE/ADSL study in 2013, we also tried to measure expectations in a more quantifiable way, i.e., we asked our participants what they expect in certain situations. We asked for more technical QoS parameters, e.g., "How fast should your internet connection at home be if you access a typical website" with answer options 1 MBit/s, 4 MBit/s etc. and we asked for more perceptional assessments, e.g. "How long should it take to download a 50 MB file via your internet connection at home" with answer options 10 seconds, 30 seconds etc. All questions and answer options *directly relates* to our test scenarios, i.e., during our study our users had to download a 50 MB file with a certain network speed and they had to evaluate it via a standard 5 point MOS scale and a binary question regarding the acceptance.

According to Table 1, most of our users stated higher expectations via the questionnaire, e.g., a download speed of 14 MBit/s was rated as non acceptable via the questionnaire but the real download of the file with this connection speed was evaluated as acceptable by the same user. For both quality indicators, MBit/s and seconds, the distribution regarding lower/equal/higher expectations is more or less consistent. Therefore, our participants were able to match both indicators.

---

[1]In [11] the authors state that the desired expectations are rather stable over time, but according to the (fast) technological development and changing usage behaviour, the adaption of desired expectations over a period of 2 years in the context of QoE seems valid.

50 MB file size (Expectations)				
	Fixed access		Mobile access	
Expectations stated via questionnaire	Speed [MBit/s]	DL duration [sec.]	Speed [MBit/s]	DL duration [sec.]
lower	11%	11%	11%	11%
equal	36%	23%	38%	26%
higher	52%	65%	50%	61%

Table 1: Percental amount of users which describe their expectations via a questionnaire in relation to real acceptance ratings.

50 MB file size (Difficulty)				
	fixed access		mobile access	
Expectations stated via questionnaire	Speed [MBit/s]	DL duration [sec.]	Speed [MBit/s]	DL duration [sec.]
lower	1	3	2	1.5
equal	2.38	2.4	1.85	2.67
higher	2.58	2.86	2.53	2.14

Table 2: Average values which describe how difficult it was for the test participants to answer the expectation questions: 1=very easy to answer, 5=very difficult

To verify how easy it was to answer the connection-speed and the download-duration questions, the test participants had to indicate how difficult it was to answer the expectations questions, see Table 2. Surprisingly, there is a small (non significant[2]) tendency, that it was more difficult for the users to state (tangible) download duration values than (abstract) download speed values. This self-evaluation shows that both types of quality indicators, abstract MBit/s & tangible seconds, are perceived in the same way by the users regarding imaginative power.

In contrast to the previously described binary acceptance question, the MOS ratings of the download scenarios are not correlated in this study with the expectations questionnaires, see Figure 5. If a correlation would exist, the MOS rating for a low downlink speed, e.g. 4 Mbit/s, would be higher for users who stated lower expectations.

3.1.4. Emergence of expectations

To examine the emergence of expectations, we asked our test participants in the LTE/ADSL study about their me-

---

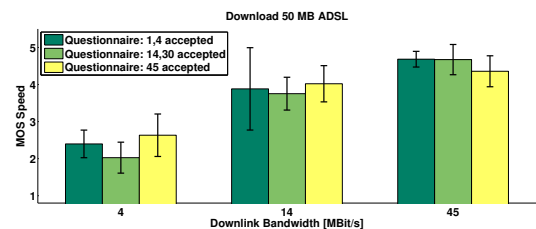[2]All confidence intervals of the mean values were overlapping.

Figure 5: Influence of stated expectations via questionnaire on MOS ratings. The overlapping CIs indicate, that there is no significant influence of indicated expectations and real rating behaviour.

dia consumption behavior and their internet usage background. There is a small, significant [3] correlation between surfing a news site via (pretended) LTE access and the consumption of LTE TV commercials: users how remembered LTE TV commercials rated more critical than users who did not remembered any TV LTE commercial. The TV ads and its promises regarding the advantages of LTE lead to higher expectations which results in lower ratings.

We also asked about media consumption to indirectly measure advertisement views. We assume that higher media, and therefore higher advertisement consumption lead to higher expectations and to lower ratings. This assumption could be proven for some scenarios, e.g., for the ADSL scenario, a higher newspaper consumption lead to more critical ratings[4]. Also for the ADSL download scenario higher radio consumption lead to more critical ratings[5]. But also the opposing has been observed: for the upload LTE scenario and the Google Maps LTE scenario, the users rated more tolerant when they also listen to radio more often[6].

3.2. Challenges of triggering explicit expectations

Provoking expectations requires an appropriate setup. In contrast to our previously described setup of provoking expectations via a manual internet connection switcher, some of our other test setups lead to more ambiguous results.

3.2.1. Influence of Video on Demand contracts

In a video study we conducted in 2011, our test users have been assigned to one of three VoD-contracts (gold, silver and bronze), differing in available movies, support levels, placement of commercials, and mainly the price charged to the user. The three contracts are presented on the TV screen and afterwards every user was randomly assigned to one contract type, which they were asked to use as reference while watching the following video snippets. There were three content classes (action, documentation, sport), each video was coded via h.264 in 1080i and with a video bitrate of 1000, 5500 and 8000kBit/s and after each consumption, the user had to evaluate the video quality via a standard MOS scale. For higher video bitrates (1000 and 5500 kBit/s), there were no clear differences regarding the assigned VoD contract, see Figure 6. There is a small, but not significant tendency, that users with a silver contract evaluate the presented video quality more critical than users with gold or bronze contract.

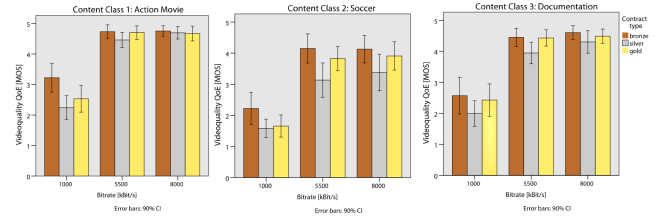It seems that the provoking attempt ("Imagine to use the VoD-contract while evaluation the presented video

---

[3]r=0.39, p <0.05
[4]r=0.54, p <0.05
[5]r=0.31, p <0.05
[6]r=0.48, p <0.05 / r=0.41, p <0.05

Figure 6: Comparison between different content classes and assigned Video on Demand contract tpyes
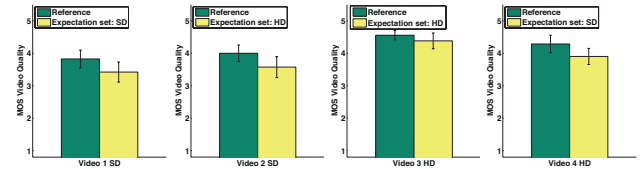

Figure 7: Influence of video quality references on perceived video quality

quality") might not be severe enough. Instead of *imagination*, a more realistic approach should be used, which could be for example a field study with a real VoD service.

3.2.2. Presented evidence

In another study, our test participants had to evaluate the presented video quality after an explicit reference regarding the approaching quality has been shown, e.g. "The following video is presented in High Definition quality (HD)". According to Fig 7, there were no significant influences of the presented hint on the video quality evaluation. Videos with a displayed explanation (e.g. "The next video will be presented in SD quality") seemed to be *always* rated a little bit more critical compared with no explanation[7].

3.3. Hidden expectations

The previously mentioned experiments tried to explicitly evoke and measure expectations. Of course, every time a quality evaluation is done, expectations are taken into account and most of the time, these expectations are unconscious, unmeasured and not verbalized. In such study setups, is it barely possible to handle expectations as independent variables and to examine their role in the whole quality evaluation process.

For example, in one of our previous studies, we examined the influence of charging and payment decision on video quality perception [13]: Our participants received a deposit of € 10, which could be used to enhance the video quality of a user-selected video. There were 20 quality

---

[7]However, this difference is statistically not significant as also evident from the overlapping confidence intervals.
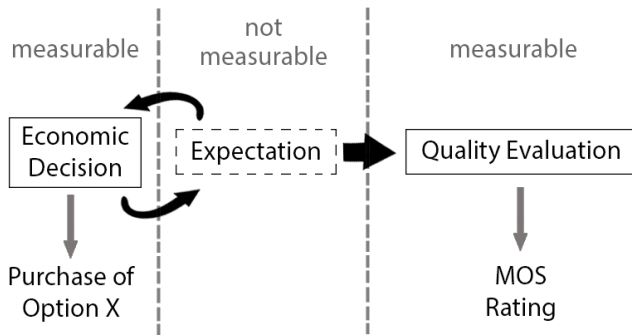
116

Figure 8: The made economic decisions and the resulting quality evaluation are measurable, but the intermediate expectation remains concealed. Decisions influence expectations and vice versa.

levels to choose from, the lowest one was always for free and the highest one was the most expensive[8]. Therefore, our users had to made a trade-off between money to spent and improved video quality. Each quality/monetary decision lead to certain expectations ("If I pay more, I expect a higher quality"), but existing expectations also influence the decision in advance. This interdependency influences quality perception and evaluation, but it is also hidden from explicit measurement, i.e., only the MOS rating as a measured result is visible. In contrast to the described disadvantages in Section 3.2, this methodological approach leads to more valid results, i.e., the expectations are set via realistic, economic decisions.

4. Conclusions and Outlook

In this paper, we presented various definitions of the term 'expectations' and different attempts to understand its nature. The discussed concepts and ideas are based on work performed in various research fields, e.g., Human Computer Interaction, Economical Studies, Social-Psychology, etc. Particularly the distinction between *desired* and *adequate* expectations and the underlying concept of acceptance perfectly fits with already existing QoE frameworks, such as the influence of user and context variables on quality perception and evaluation. Also, expectations are not static, but change over time and are strongly influenced by contextual variables.

To include expectations as an influencing factor in QoE research, the *emergence*, the *triggering* and the *measurement* of expectations have to be considered and properly accounted for. As exemplified by the presented mobile vs. fixed Internet access study in Section 3.1, user background related factors such as the usage of media, the everyday consumption of advertisements and the individual Internet usage history can be used to understand the emergence of certain expectations. The deliberate

triggering of specific expectations requires some effort: the VoD and HD/SD studies presented in Section 3.2 demonstrate that realistic settings are needed to evoke expectations in a valid fashion. Our studies show, that user are able to verbalize specific expectations regarding the download duration of a file or the desired download speed.

Measuring and explicitly triggering of expectations might not be feasible for all QoE-related studies since expecations are fragile: they cannot be assessed without also influencing them, analogous to the Schrdinger's cat problem quantum physics. Therefore, we propose to pay more attention to factors influencing emergence of expectations (media, advertisements, internet usage, etc.) rather than gauging expectations directly. The reason is that getting insights into the emergence of expectations is more unobtrusive and less interfering than explicitly triggering and measuring expectations by e.g. direct interrogation of participants.

As regards future work, the expectations framework presented in Section 2 already includes binary acceptance in relation to desired and adequate expectations. However, it is still unclear how more fine-grained opinion scores (e.g. MOS) can be integrated in this framework in order to better relate expectation to quality evaluation.

5. References

[1] P. Le Callet, S. Möller, and A. Perkis, "Qualinet white paper on definitions of quality of experience (2012)," European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003), Tech. Rep., 2012, Lausanne, Switzerland.

[2] I. T. Union, "Vocabulary and effects of transmission parameters on customer opinion of transmission quality, amendment 2," ITU-T Recommendation P.10/G.100, Tech. Rep., 2006.

[3] R. Schatz, T. Hoßfeld, L. Janowski, and S. Egger, "From packets to people: Quality of experience as a new measurement challenge," in *Data Traffic Monitoring and Analysis*, ser. Lecture Notes in Computer Science, E. Biersack, C. Callegari, and M. Matijasevic, Eds. Springer Berlin Heidelberg, 2013, vol. 7754, pp. 219–263.

[4] J. A. Bonito, J. K. Burgoon, and B. Bengtsson, "The role of expectations in human-computer interaction," in *Proceedings of the international ACM SIGGROUP conference on Supporting group work*, ser. GROUP '99. New York, NY, USA: ACM, 1999, pp. 229–238.

[5] J. K. Burgoon and S. B. Jones, "Toward a theory of personal space expectations and their violations," *Human Communication Research*, vol. 2, no. 2, pp. 131–146, 1976. [Online]. Available: http://dx.doi.org/10.1111/j.1468-2958.1976.tb00706.x

[6] R. T. Watson, L. F. Pitt, and C. B. Kavan, "Measuring information systems service quality: Lessons from two longitudinal case studies," *MIS Quarterly*, vol. 22, no. 1, pp. 61–79, 1998.

---

[8]In the study the actual price depended on the set price plan.

[7] A. Parasuraman, V. A. Zeithaml, and L. L. Berry, "SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality," *Journal of Retailing*, vol. 64, no. 1, pp. 12–40, 1988. [Online]. Available: http://search.ebscohost.com/login.aspx?direct=true\&\#38;db=buh\&\#38;AN=6353260\&\#38;site=ehost-live

[8] D. J. Kim, D. L. Ferrin, and H. R. Rao, "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*, ser. ICEC '03.   New York, USA: ACM, 2003, pp. 310–315.

[9] K. Kristensen, A. Martensen, and L. Gronholdt, "Measuring the impact of buying behaviour on customer satisfaction," *Total Quality Management*, vol. 10, no. 4-5, pp. 602–614, 1999. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/0954412997587

[10] Y.-H. Hsieh and S.-T. Yuan, "A quantitative approach to measure customer expectation for service innovation within service experience delivery," in *Proceedings of the 11th International Conference on Electronic Commerce*, ser. ICEC '09.   New York, NY, USA: ACM, 2009, pp. 38–46.

[11] V. Zeithaml, L. Berry, and A. Parasuraman, "The nature and determinants of customer expectations of service," *Journal of the Academy of Marketing Science*, vol. 21, no. 1, pp. 1–12, 1993. [Online]. Available: http://dx.doi.org/10.1177/0092070393211001

[12] A. Sackl, K. Masuch, S. Egger, and R. Schatz, "Wireless vs. wireline shootout: Howuser expectations influence quality of experience," in *Quality of Multimedia Experience (QoMEX), 2012 Fourth International Workshop on*, 2012, pp. 148–149.

[13] A. Sackl, P. Zwickl, S. Egger, and P. Reichl, "The role of cognitive dissonance for qoe evaluation of multimedia services," in *Globecom Workshops (GC Wkshps), 2012 IEEE*, Dec., pp. 1352–1356.