

Alone or Together: Measuring Users' Viewing Experience in Different Social Contexts

Zhu Yi^{1a}, Ingrid Heynderickx^b, Judith A. Redi^a

^aDelft University of Technology, Mekelweg 4, Delft, The Netherlands, 2628 CD

^bEindhoven University of Technology, Postbus 513, Eindhoven, The Netherlands, 5600 MB

ABSTRACT

In the past decades, a lot of effort has been invested in predicting the users' Quality of Visual Experience (QoVE) in order to optimize online video delivery. So far, the objective approaches to measure QoVE have been mainly based on an estimation of the visibility of artifacts generated by signal impairments at the moment of delivery and on a prediction of how annoying these artifacts are to the end user. Recently, it has been shown that other aspects, such as user interest or viewing context, also have a crucial influence on QoVE. Social context is one of these aspects, but it has been poorly investigated in relation to QoVE so far. In this paper, we report the outcomes of an experiment that aims at unveiling the role that social context, and in particular co-located co-viewing, plays within the visual experience and the annoyance of coding artifacts. The results show that social context significantly influences user's QoVE, whereas the appearance of artifacts doesn't have impact on viewing experience, although users can still notice them. The results suggest that quantifying the impact of social context on user experience is of major importance to accurately predict QoVE towards video delivery optimization.

Keywords: Social Context, Quality of Viewing Experience, Visual Quality

1. INTRODUCTION

Billions of digital videos cross IP networks every day, and this number is only destined to grow in the coming years [1]. As a consequence of technological limitations (e.g., bandwidth and storage constraints) most uploaded videos are encoded, which may result in visible compression artifacts, such as blockiness, ringing and blur, in the decoded video [2]. Such artifacts may severely degrade perceived quality which, in turn, may influence end-user's satisfaction with the video service. Moreover, evidence shows that there is a positive correlation between the willingness to pay for a multimedia service/system and the quality of the video offered to the user [3]. Hence, measuring and predicting user's satisfaction becomes very important for stakeholders in the online multimedia delivery business.

User's satisfaction with an online video can be measured with the concept Quality of Visual Experience (QoVE), reflecting the degree of delight or annoyance of the user with a video [4]. Objective QoVE estimation for online video has so far been based on two types of measurements: network-based and signal-based [5]. Network-based measurements aim at predicting QoVE based on information gathered from network and service conditions without accessing the decoded video [6]. Such information contains parameters such as e.g., packet loss ratio and network delay, which are also referred to as network Quality of Service (QoS) parameters [7]. Video-related network QoS parameters, such as buffering time and buffer ratio are also often considered [8]. Signal-based measurements, on the other hand, estimate the visibility of artifacts generated in the video throughout its lifecycle by means of signal analysis techniques. Artifact visibility measures are therefore performed on the decoded bit stream and are linked to an overall measure of QoVE for the end user [9-10]. Although achieving remarkable results, both network- and signal-based measurements mainly address perceptual aspects of the viewing experience, and often fail to capture the overall user satisfaction with the multimedia delivery system [4].

¹ *Y.Zhu-1@tudelft.nl; phone: + 31 (0)15 2782549

Lately, evidence has been provided that other aspects, e.g., user interest [11] and personality [12] also have a crucial influence on QoVE. In particular QoVE is described as the result of the interaction of a set of factors, not necessarily independent and not limited to either QoS parameters nor visibility of artifacts [4, 9]. In [4], three classes of factors are suggested to influence QoVE. System factors usually refer to properties that determine the technically produced quality of online video. As such, they include QoS parameters, video genre, characteristics of the device on which they are displayed and the signal/media configuration [4]. Human factors are related to the user's personal characteristics, such as age, gender, personality and interest. Finally, contextual factors describe all aspects of the user's environment, e.g., social interaction at the time the video is viewed, or the physical location in which the video is viewed. As mentioned earlier, little research has been devoted to factors other than the system ones, and within the contextual factors, especially social context has been poorly considered so far.

Social context usually refers to the relationships and interactions that occur between the user and other people (being them family, friends or strangers) during the experience [13]. In the past, many studies have reported the social impact of traditional TV watching [14,15]. TV viewers react emotionally to TV programs: they record and share their favorite programs with families and friends, and they discuss the program either during the show or even afterwards. As such, TV programs have provided a kind of "social glue" that bonds friends and strangers together by letting them have some common topic to chat about [16]. Talking about TV programs (e.g., yesterday's football match or a popular TV show) is helpful for people to start and sustain relationships with each other [17]. Furthermore, it has been shown that co-located co-viewing is a rather common way of consuming media, e.g., watching World Cup soccer games with friends on a big screen [18], and that co-viewing when enjoying each other's company can increase user's overall satisfaction [19]. However, the direct link between social context and eventual QoVE still remains unexplored.

This paper aims at understanding how the social context influences QoVE when watching infotainment videos. By social context we specifically indicate here the presence/absence of a group of known people (friends) sharing a viewing experience while being in physical proximity. We report the outcomes of an experiment involving two (disjoint) groups of participants, viewing a set of videos in two different social situations. Participants in the first group watched the videos by themselves, whereas participants in the second group watched the same set of videos along with two friends. Videos were compressed at two different bitrates, one of which resulting in clearly visible compression artifacts. After viewing each video, participants judged QoVE on five aspects, namely Enjoyment [20], Satisfaction [20], Involvement [21], Endurability [22], and Perceived Visual Quality [23]. The overall goal of the experiment was to determine whether the type of social viewing situation would matter for the overall appreciation of the experience, and whether this impact was conditioned by the presence of compression artifacts in the displayed videos. Results indicate that in fact co-located co-viewing with friends enhances the overall QoVE, even when videos are compressed at low bitrates.

The remainder of the paper is organized as follows. In section 2, we present the experimental details, including its settings, stimuli, measurements, etc. Then, in section 3, we report some initial descriptive statistics of the experimental data and explain how we analyze them. The results are shown in section 4, followed by our conclusions in section 5.

2. METHODOLOGY

2.1 Experimental setup

We created two different real-life social situations, as shown in Figure 1: one situation with a single person (S) watching a video alone, and a second situation with a group of friends (G) watching a video together. Participants were seated on a couch in front of a screen (i.e., a 41" LCD-TV) in a setting that replicates common home-viewing modalities. The number of people co-viewing videos was limited to three at a time; within groups, each participant specified that he/she held a friendship relationship with the other group members.

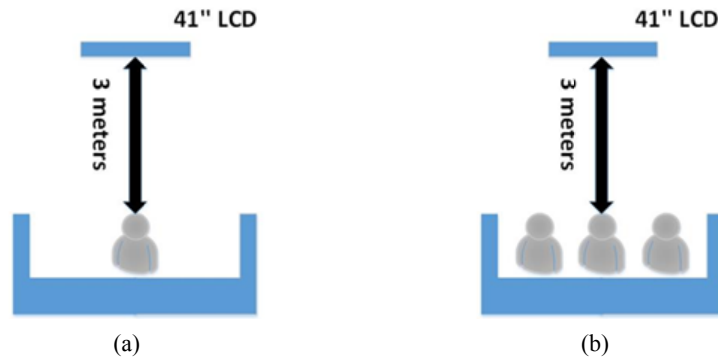


Figure 1. The two different social situations investigated in the experiment. In social situation (a), a single participant watched videos on a 41" LCD-TV from a couch 3 meters away. In social situation (b), three participants, being friends, watched the same videos in the same environmental conditions.

2.2 Stimuli

Three different genres of video, i.e., (1) *Sports*; (2) *Comedy* and (3) *Education*, were used in this experiment. To select these genres, we first conducted a pilot survey to understand which was the preferred social context for watching 15 of the most common genres of online video. For each genre, participants were asked in which social situation they would prefer to watch it: alone, or with friends. If they were not sure, they could choose “it doesn’t matter”. 80 students from TU Delft participated to the survey. A clear preference (gathering over 50% of the choices) was found for two genres, as shown in Figure 2: 51% of the students indicated a preference for watching *Comedy* videos with friends and 61% of the students indicated to prefer watching *Education* videos alone. Thus, we decided to include these two genres in this study. In addition, we also selected a genre without a clear consensus, namely *Sports*. For this genre 44% of the students indicated a preference for watching it with friends, while 46% thought the viewing situation did not matter.

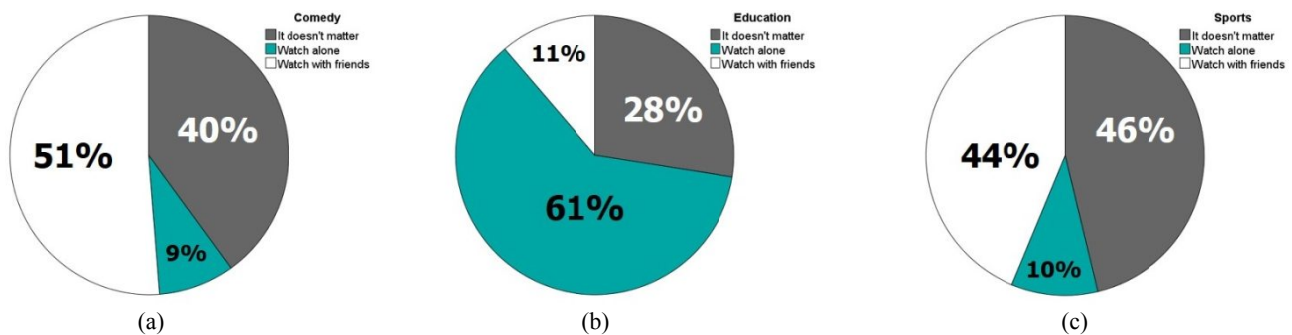


Figure 2. Results of the online survey on preferred social context for watching (a) comedy, (b) education and (c) sports videos. The white area in the pie plots indicates preference for watching alone, the greenish area indicates preference for watching with friends, and the grey area indicates that it doesn’t matter.

As a result, two different video clips were selected from each genre (screenshots are given in Figure 3): an excerpt of the Jimmy Kimmel Show (JKS) and Saturday Night Live (SNL) for comedy, a documentary on Birds of Paradise (BoP) and a TED talk (TED) for education, and a fragment of a Soccer and a Basketball game for sports.

All videos were downloaded from YouTube, encoded in H.264/AVC, which is the most commonly used format for online videos [24]. The videos were at least 5 minutes long with bitrates varying between 2955 kbps and 1737 kbps. All the original videos were 30 frame per second (fps) with a spatial resolution of 1280x720 pixels. All videos were further encoded in H.264/AVC at two different bitrates: high (2000 kbps) and low (600 kbps). The reason to choose only two bitrate levels was that we were not interested at this stage in the changes of user QoVE generated by a change in bitrate;

rather, we intended to focus on the changes in QoVE with respect to presence/absence of co-viewers, given a certain bitrate level. It should also be pointed out that the original bitrates of the BoP and TED video were less than 2000 kbps. So for these two videos, we used the original video bitrate as the high bitrate setting. Eventually, low bitrate versions of the videos presented clearly visible compression artifacts, whereas this was not the case for the high bitrate versions. Finally, the audio of all clips was encoded in the AAC format with a (high) bitrate of 112 kbps to limit the impact of audio impairments on user judgments.

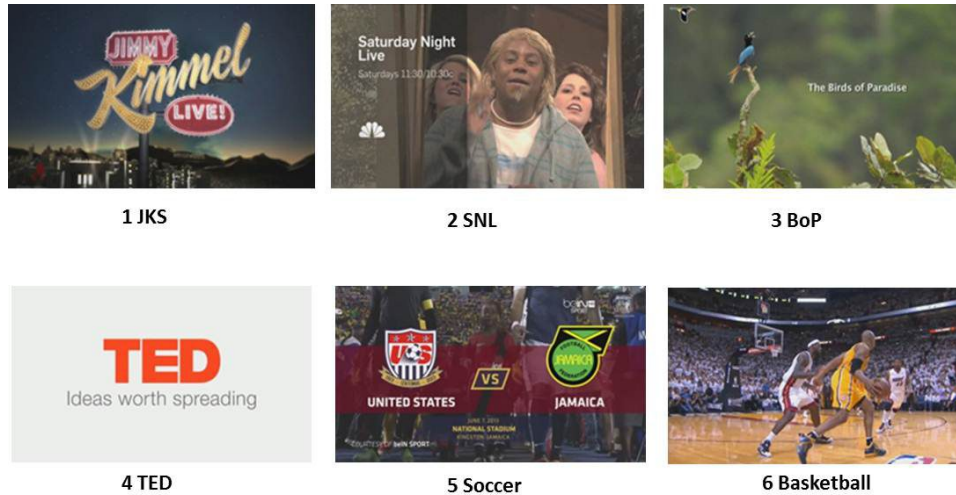


Figure 3. Screenshots of the six video clips used in this study

In addition, three 10s-long video samples, one per genre, were selected for making the participants acquainted with the experimental procedure and questionnaires before the start of the actual experiment. The media configuration of these samples is the same as for the test videos (30 fps, 1280x720, H.264/AVC). The samples were also encoded at the same bitrate levels (i.e., 2000 kbps and 600 kbps).

2.3 Participants

Sixty participants (i.e., 27 females and 33 males; 30 per viewing condition) were recruited for this experiment. Most participants were students or employees of the Delft University of Technology and aged between 18 and 41 years (mean age = 26.5). Participants involved in one type of social situation (e.g., single viewing) did not see videos in the other situation (e.g., group viewing). Participants within each social situation were further divided into two sub-groups, as shown in Table 1, each of which watched the same video content only once at a certain bitrate level (i.e., either high or low). Every participant saw all 6 different video contents once during the experiment, thus watching two videos per genre. As a result, our experiment had a mixed design, investigating the effect of social situation and bitrate level between-subjects, whereas the effect of video genre was investigated within-subjects.

Table 1 Overview of the experimental setup. V indicates the test Video Clip, whereas S indicates a group of participants that watched the videos in a single viewing condition and G indicates the group of participants that watched the videos with multiple people

	Genre 1		Genre 2		Genre 3	
	V1	V2	V3	V4	V5	V6
High Bitrate	S1 G1	S2 G2	S1 G1	S2 G2	S1 G1	S2 G2
Low Bitrate	S2 G2	S1 G1	S2 G2	S1 G1	S2 G2	S1 G1

2.4 Apparatus

All stimuli were visualized on a 41" LCD TV. All of the TV's parameters, such as brightness and contrast, were set to their default value for the experiment. The viewing distance was 6 times the height of the screen (approximately 3 meters) in order to satisfy the preferred viewing distance [23] (see also figure 1). The rest of the environmental settings followed the ITU-R BT.500-13 Recommendation [23] and were kept the same for all participants. Also the audio volume was kept constant for all participants. Participants used a tablet (Samsung Galaxy Tab 2) to fill in all the questionnaires.

2.5 Measurements

Since there is still no clear consensus on how to measure QoVE subjectively, we considered 5 different aspects to be important. These aspects included satisfaction, involvement, enjoyment and endurance of the visual experience, as well as perceived visual quality (PVQ). Satisfaction measures the user's overall feelings about the viewing experience, and as such is considered relevant for QoVE. To measure satisfaction we selected the 4 items of the questionnaire used in [20]. Some studies suggested that the QoVE measurement should be complemented by a measure of the level of enjoyment, which reflects how much pleasure a user gets from experiencing a video [25]. To measure user enjoyment, we used again 4 items from the questionnaire in [20]. Involvement, which refers to user awareness to the physical environment around him/her while living the viewing experience, is a concept mainly known from virtual reality media delivery, but as such may also reflect user's QoVE. Indeed, one study showed that the more a user is involved in a service, the better the experience will be judged [26]. Therefore, we measured involvement through an adapted version of the Igroup Presence Questionnaire (IPQ) [21], the latter being mainly known from presence research in virtual reality environments. Also here we limited the measurement of involvement to 4 items. The aspect of "endurability" was added to the questionnaire, since it has been used to describe the consequence of high QoVE: a highly appreciated experience increases the likelihood of it being remembered, as well as the user intention to repeat it, share it with friends or recommend it [26]. As such, we consider the willingness to repeat the visual experience ("endurability") as an indicator of high QoVE, and to measure it we used four adapted items from O'Brien's questionnaire [22]. As shown in appendix, the above four aspects were all measured on seven-point Likert scales (ranging from "fully disagree" to "fully agree"). To complement the questionnaire with the more traditional aspect of quality of experience related to artifact visibility, participants were also asked to indicate their perceived level of perceptual visual quality (hereafter referred to as PVQ) on two five-point scales: an overall quality scale, and an impairment annoyance scale according to [23]. Eventually, the questionnaire consisted of 18 items.

2.6 Procedure

Firstly, participants were seated on a couch in front of the 41" LCD TV and filled out a demographic questionnaire. Before the start of the actual experiment, an introduction was given to all the participants. In this introduction, we showed six 10-second video samples (spanning a broad range of visual quality) and a questionnaire sample to make them acquainted with the range of visual quality and with the questionnaire scoring scales. After that, participants watched the six experimental video clips in a random order, and with alternate quality level (according to Table 1). Following the presentation of each video clip, participants were asked to fill out the questionnaire described in section 2.5. Items in the questionnaire were randomized differently for each participant. The next clip would not be played until all participants in a session completed the questionnaire. In the group viewing situation, one tablet was given to each participant, and they were asked to fill in the questionnaire individually, i.e., they were not allowed to discuss their impressions and judgments while compiling the questionnaire. They were instead allowed to interact while watching the video clips.

3. DATA PREPARATION

The experiment outputted Aspect Scores (AS) in four aspects (i.e., satisfaction, involvement, enjoyment and endurance) on a 7-point Likert scale and a set of Opinion Scores (OS) for PVQ on a 5-point scale.

We define here the Aspect Score $AS(v_n^{b,s}, k_a, p)$ as the score given by participant p ($p = 1, \dots, P$) to item k_a , $k=1, \dots, 4$, in aspect $a \in \{\text{satisfaction, involvement, enjoyment, endurance}\}$ for video $v_n^{b,s}$. Here, $v_n^{b,s}$ represents the video v_n , $n = 1, \dots, 6$ of those listed in section 2.2, at a bitrate level b , $b \in \{2000 \text{ kbps}, 600 \text{ kbps}\}$ in social situation of type s , $s \in \{\text{single, group}\}$. The two items measuring PVQ output the Opinion Score $OS(v_n^{b,s}, i_{PVQ}, p)$, with a notation similar to the one presented above and $i_{PVQ} \in \{1 = \text{overall quality}, 2 = \text{annoyance}\}$.

To prepare the data for the analysis, we first calculated the Cronbach's Alpha (α) among items in each aspect, the results of which are shown in Table 2. Cronbach's alpha reflects internal consistency between the items measuring a given aspect, and as such, is used to assess the internal reliability of a questionnaire [27]. Values of α above 0.8 are considered to represent high consistency, i.e., the different items measure the same, underlying psychological construct. For four out of the five aspects of QoVE that we measured, the Cronbach's alpha was considerably higher than 0.8, indicating that the different items in each aspect questionnaire measured a consistent underlying concept. As a consequence, we could sum the scores on the individual items per aspect to obtain an overall Aspect Score:

$$AS(v_n^{b,s}, p) = \sum_{k_a=1}^4 AS(v_n^{b,s}, k_a, p) \quad a \in \{satisfaction, enjoyment, involvement, durability\} \quad (1)$$

As a result, for these four aspects, the Aspect Scores varied between 4 and 28 (resulting from 4 items on a 7-point scale). The exception is the Cronbach's alpha of PVQ, indicating that the two questions related to it measure a different concept. Indeed, the Spearman correlation between the scores given to the two PVQ items was only 0.632, suggesting that there is some relation between overall perceived quality and artifact annoyance, but that nonetheless overall perceived quality measures (most probably) a broader concept than just artifact visibility. As a consequence, we analyzed the two items of PVQ separately in the remainder of our paper.

Table 2 Internal Consistency measured by Cronbach's alpha among items in each aspect of the questionnaire

Aspect	Enjoyment	Involvement	Satisfaction	Endurability	PVQ
Chronbach's α	0.917	0.82	0.961	0.939	0.773

4. RESULTS

Because of our mixed experimental design, we examined all data (i.e., AS of all aspects and OS of PVQ) using an ANOVA with repeated measures (by means of the SPSS software version 19). Bitrate level b and social situation s were set as between-subjects factors, whereas video genre (i.e., comedy, education or sports) was investigated as within subject factor. The aggregate scores for durability, enjoyment, involvement and satisfaction, as well as the OS for the PVQ items were the dependent variables, and the ANOVA model also included all 2-way interactions among the three factors (i.e., social situation, bitrate level and video genre).

This analysis showed that the type of social situation (i.e., single or group) significantly influenced user's QoVE in terms of enjoyment ($F(1,116) = 4.518, p = 0.036$) and durability ($F(1,116) = 3.855, p=0.052$). We didn't find a significant effect of social situation on satisfaction ($F(1,116) = 1.563, p = 0.214$), involvement ($F(1,116) = 2.011, p = 0.159$) and perceived quality (for PVQ1: $F(1,116)=0.113, p=0.738$ and for PVQ2: $F(1,116)=0.554, p=0.458$).

Related to enjoyment, figure 4(a) shows that participants who watched videos with friends rated their enjoyment higher than those who watched the same videos alone. This finding confirms what was already found in previous studies on traditional TV [18], i.e., users love watching videos with friends, and such social interaction increases their level of enjoyment. We didn't find a significant interaction between social situation and genre ($F(2, 232) = 0.898, p = 0.409$), suggesting that a higher enjoyment when watching a video in group instead of alone is independent of the genre of the video (at least for the genres investigated in this experiment). Figure 4(a) though suggests that the difference in enjoyment between watching alone and watching in group is slightly higher for the comedy videos than for the other two genres, which would be in line with the results of our pilot online survey that people prefer watching comedy videos with friends (see section 2.2).

In terms of durability, figure 4(b) shows that participants that watched the videos in group gave a higher score than those who watched the videos on their own. This finding indicates that when participants watched the videos with friends, they were more willing to recommend videos to others and more eager to repeat the viewing experience. Again, we didn't find a significant interaction between social situation and genre ($F(2,232) = 0.412, p = 0.663$), which indicates that user's durability in different social situations was independent of the genre of the video. Nevertheless, figure 4(b) also suggests that the difference in durability between two social situations was more pronounced for comedy videos than for the other two genres.

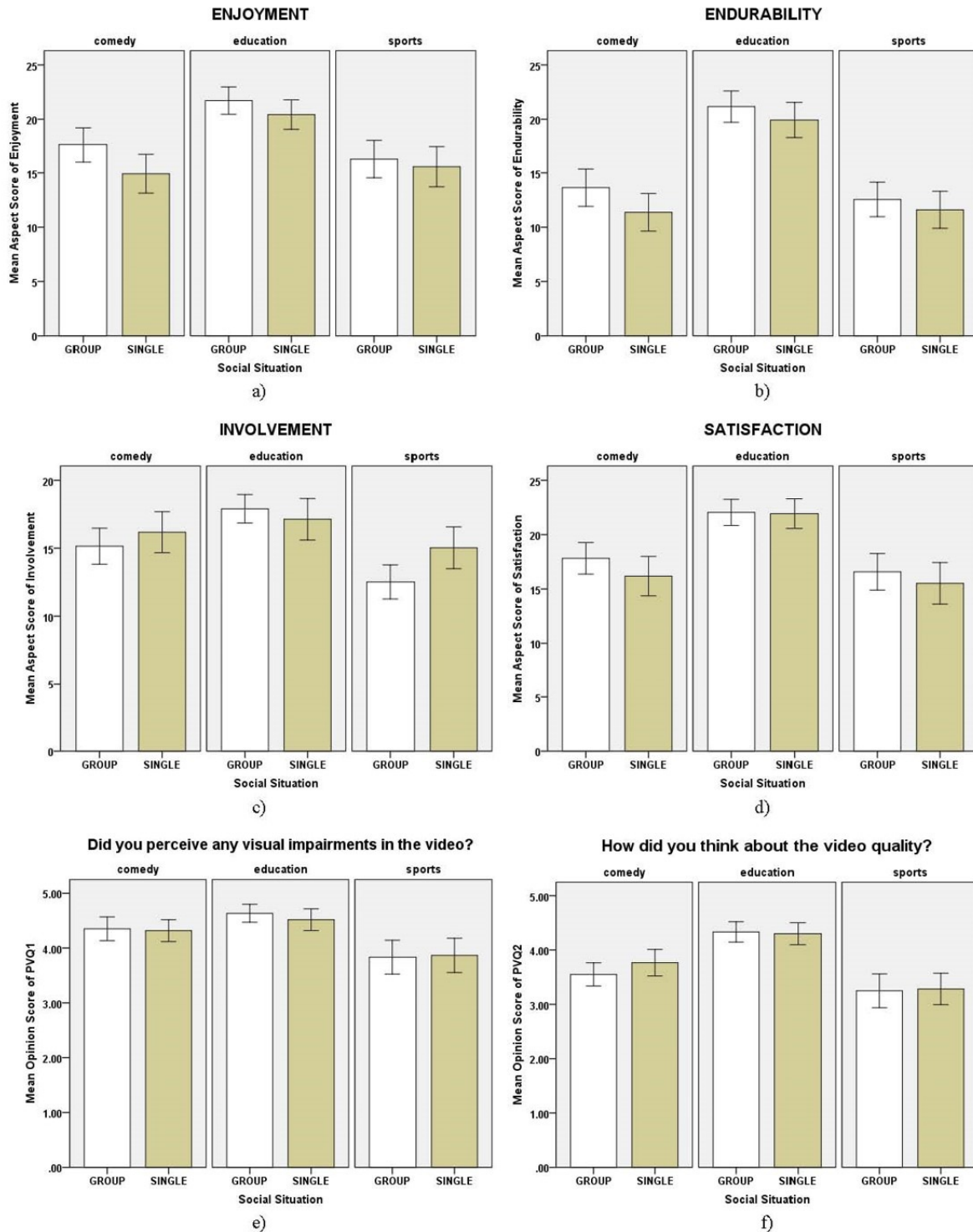


Figure 4. The mean AS and OS between two social situations among three genres. Histogram (a) represents the mean AS of enjoyment; histogram (b) represents the mean AS of endurance; histogram (c) represents the mean AS of involvement; histograms (d) represent mean AS of satisfaction and histograms (e, f) represent mean Opinion Score of two PVQ items. The white bars give the score for group viewing while the yellow bars give the score for single viewing. The scores for the three genres (i.e., comedy, education and sports) are shown in three separated columns.

We didn't find a significant effect of social situation on involvement, but we did find a significant interaction between social situation and genre ($F(2, 232) = 3.408, p = 0.035$), suggesting that user's involvement in different social situations varies with video genre (see also figure 4(c)). In order to better understand the effect of video genre on user's involvement, we split up the data file per video genre, and repeated the ANOVA (but now with only social situation and bitrate as the independent factors, including their interaction). The results showed that people who watched the sports videos with friends were less involved in the video than those who watched these videos alone ($F(1,116) = 6.363, p = 0.013$). No significant difference in involvement between the two social situations was found for the education and comedy videos ($F(1,116) = 0.683, p = 0.410$ and $F(1, 116) = 1.053, p = 0.307$ respectively). The latter finding may be explained by the fact that education and comedy videos include storytelling, which needs to be followed for the experience to be meaningful. This is not true for sports videos, and as a consequence users may be more open to social interaction when watching those.

Finally, we didn't find a significant effect of social situation on satisfaction, and also the interaction between social situation and video genre was not significant for the aspect of satisfaction ($F(2,232) = 0.557, p=0.574$), as can be deduced also from figure 4(d).

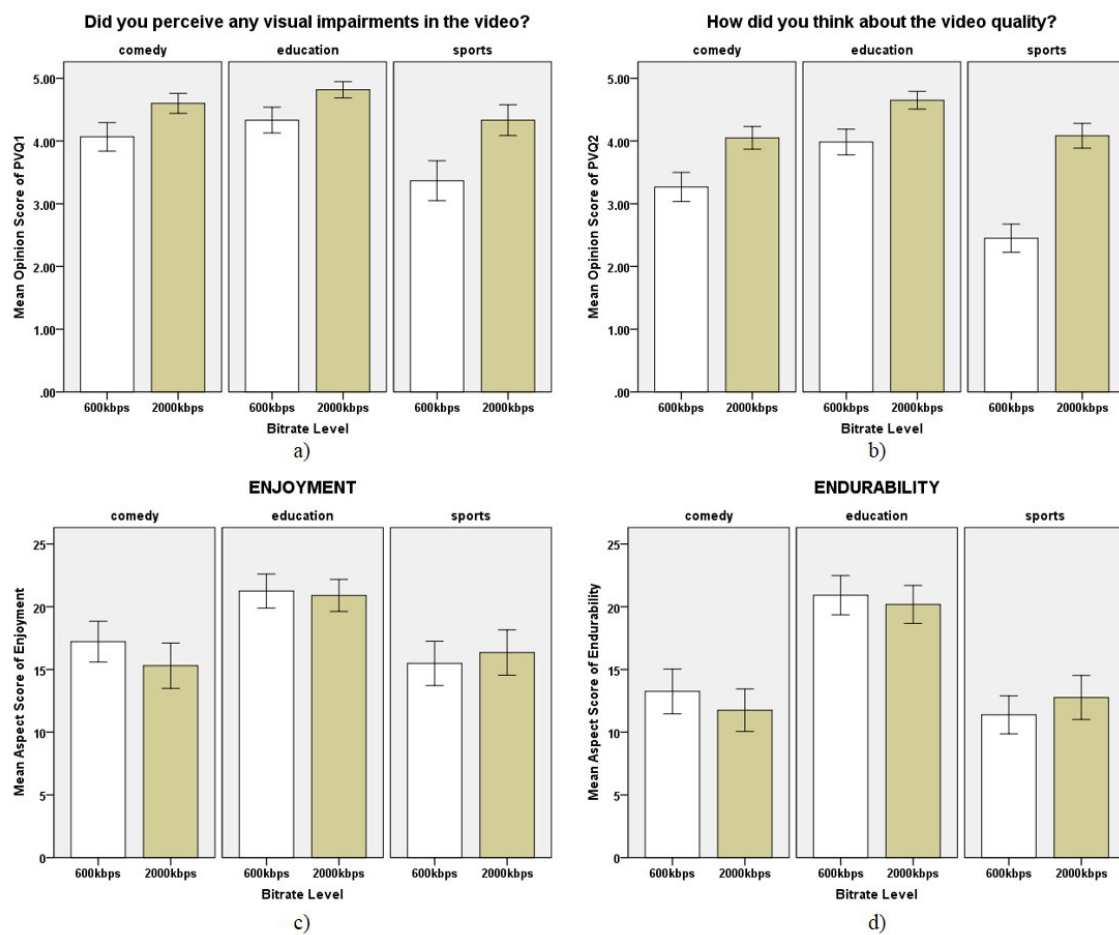


Figure 5. The mean OS and the mean AS between two bitrate levels (600kbps and 2000kbps) among three genres. Here, the histogram (a, b) represent the mean OS of two items between two bitrate levels; the histogram (c, d) represent the mean AS of enjoyment and endurability between two bitrate levels respectively. The white bars indicate the low bitrate level while the yellow bars indicate the high bitrate level. The mean AS among three genres (i.e., comedy, education and sports) are shown in three separated columns.

As mentioned above, the PVQ ratings were not affected by the social situation, suggesting that the participants' tolerance to visual artifacts in the video did not change with the presence of co-viewers (shown in figure 4(e, f)). We did though find a significant effect of bitrate level on both PVQ items (for PVQ1: $F(1,116)=32.523, p<0.001$ and for PVQ2:

$F(1,116)=112.162, p<0.001$). Figure 5 (a, b) shows that participants scored the overall quality and (the inverse of) the annoyance of low bitrate videos significantly lower than those of high bitrate videos. On the other hand, we found no significant effect of bitrate level on enjoyment ($F(1,116) = 0.407, p = 0.525$, figure 5(c)), satisfaction ($F(1,116) = 0.101, p = 0.751$), endurance ($F(1,116) = 0.141, p = 0.708$, figure 5(d)) and involvement ($F(1,116) = 0.363, p = 0.548$). Hence, in combination, these results suggest that participants could notice the presence of artifacts in the videos, as well as indicating their annoyance and impact on the overall visual quality; nevertheless, this did not affect the overall appreciation for the experience with the video. In addition, no interaction effect was found in any aspect between bitrate level and social situation, suggesting that the difference in user's QoVE between the two social situations was independent of the bitrate level of the videos.

Finally, we also computed the Spearman correlation coefficients between the *AS* of enjoyment, satisfaction, endurance and involvement and the *OS* of two PVQ items. As shown in table 4, four *AS* aspects (i.e., enjoyment, satisfaction, endurance and involvement) were highly correlated with each other, but poorly correlated with the two PVQ items (all <0.4), suggesting that PVQ can't fully represent user's overall appreciation of the experience. Such results further illustrate that if we are interested in delighting users and letting them repeat the viewing experience (which is in the core interest of the display manufacturer or video service provider), then we should start caring about other things beyond PVQ and understand how factors such as social context or genre can compensate for low visual quality.

Table 4. The Pearson correlations between *AS* of Enjoyment, satisfaction, endurance and *OS* of two PVQ items.

	PVQ2	Enjoyment	Satisfaction	Endurability	Involvement
PVQ1	0.632	0.229	0.315	0.245	0.190
PVQ2	—	0.290	0.386	0.351	0.258
Enjoyment		—	0.902	0.859	0.618
Satisfaction			—	0.861	0.643
Endurability				—	0.580

5. CONCLUSIONS

In this paper, we investigated the impact of social context (i.e., physical presence of co-viewers) on the Quality of the Viewing Experience. Twelve videos, encoded at different bitrates and belonging to different genres, were watched by two disjoint groups of participants, who watched the videos alone, or with two of their friends, respectively. QoVE was then measured through a questionnaire investigating 5 aspects: enjoyment, endurance, satisfaction, involvement in the viewing experience and perceived visual quality. The results showed that co-viewing videos with friends increased participants' level of enjoyment and enhanced the endurance of the experience. On the other hand, the presence of co-viewers was shown to decrease participants' involvement with sports videos. Low bitrates were shown to have a negative effect on perceived visual quality, but they did not affect any of the other aspects of the visual experience.

These preliminary results indicate that user's QoVE does not purely depend on video bitrate, nor entirely on the annoyance that visual artifacts induced by low bitrate bring about. Social context has a meaningful impact on user enjoyment as well as on willingness to repeat the viewing experience, which is of main importance for the optimization of future multimedia delivery systems. As a result, besides perceived visual quality itself, information on the social context of the viewing experience should also be taken into account in video delivery optimization.

As a preliminary study, our current work also has limitations, e.g., limited number of genres, participants and types of social context. In the future, it will be interesting to investigate social context in more detail, e.g., looking at the impact of social relationship (i.e., friends versus strangers) or co-viewers personality on viewing experience appreciation. More work is needed also towards our ultimate goal, i.e., to quantify social context information in an automated way and incorporate such information into quality metrics in order to enable more effective optimization of media delivery systems.

REFERENCES

- [1] Cisco Visual Networking Index: Forecast and Methodology, 2010–2015, available at: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360.pdf
- [2] Chikkerur, Shyamprasad, Chikkerur, S., Sundaram, V., Reisslein, M., & Karam, L. J. "Objective video quality assessment methods: A classification, review, and performance comparison." *Broadcasting, IEEE Transactions on* 57(2), 165-182 (2011).
- [3] K. Yamori and Y. Tanaka, "Relation between willingness to pay and guaranteed minimum bandwidth in multiple-priority service," in *Proc. Joint Conf. 10th Asia-Pacific Conf. Commun. and 5th Int. Symp. Multi-Dimensional Mobile Commun.* vol. 1, 113–117 (2004)
- [4] Le Callet, P., S. Möller, A. Perkis. "Qualinet white paper on definitions of quality of experience." 2012, available at: http://www.qualinet.eu/images/stories/Qualinet_Whitepaper_v1.2.pdf.
- [5] Serral-Gracià, R., Cerqueira, E., Curado, M., Yannuzzi, M., Monteiro, E., & Masip-Bruin, X. "An overview of quality of experience measurement challenges for video applications in IP networks." *Wired/Wireless Internet Communications*. Springer Berlin Heidelberg, 252-263 (2010).
- [6] Lin, T.L., Kanumuri, S., Zhi, Y., Poole, D., Cosman, P.C., Reibman, A.R., "A Versatile Model for Packet Loss Visibility and its Application to Packet Prioritization." *IEEE Transactions on Image Processing* 19(3), 722–735 (2010).
- [7] J. Asghar, F. Le Faucheur, and I. Hood, "Preserving video quality in IPTV networks," *IEEE Trans. Broadcast.*, 55(2), 386–395 (2009).
- [8] Dobrian, F., Sekar, V., Awan, A., Stoica, I., Joseph, D. A., Ganjam, A., & Zhang, H. "Understanding the impact of video quality on user engagement." *SIGCOMM-Computer Communication Review* 41(4), 362 (2011).
- [9] Redi, Judith A. "Visual quality beyond artifact visibility." *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics (2013).
- [10] S. S. Hemami, A. R. Reibman. "No-reference image and video quality estimation: Applications and human-motivated design." *Signal Processing: Image Communication*, 25(7), 469-481 (2010)
- [11] P. Kortum, M. Sullivan. "The effect of content desirability on subjective video quality ratings." *Human factors: the journal of the human factors and ergonomics society* 52(1), 105-118 (2010).
- [12] Wechsung, Schulz, Engelbrecht, Niemann, Möller. "All Users Are (Not) Equal-The Influence of User Characteristics on Perceived Quality, Modality Choice and Performance." In *Proc. Paralinguistic Information and its Integration in Spoken Dialogue Systems Workshop*. Springer New York (2011).
- [13] Scheinkman, J. A. "Social interactions." *The New Palgrave Dictionary of Economics* 2 (2008).
- [14] Gauntlett, D., & Hill, A. "TV living: Television, culture and everyday life." New York: Routledge (2009).
- [15] Kubey, R., & Csikszentmihalyi, M. "Television and the quality of life: How viewing shapes everyday experiences." Hillsdale, NJ: Lawrence Erlbaum Associates (1990).
- [16] Lee, B., & Lee, R. S. "How and why people watch TV: Implications for the future of interactive television." *Journal of Advertising Research*, 35(6), 9–18 (1995).
- [17] Chorianoopoulos, Konstantinos, and George Lekakos. "Introduction to social TV: Enhancing the shared experience with interactive TV." *Intl. Journal of Human-Computer Interaction* 24(2), 113-120 (2008).
- [18] Morrison, M., Krugman, D. "A look at mass and computer mediated technologies: Understanding the roles of television and computers in the home." *Journal of Broadcasting & Electronic Media* 45(1), 135-161 (2001).
- [19] Oehlberg, Ducheneaut, Thornton, Moore, Nickell. "Social TV: Designing for distributed, sociable television viewing." *Proc. EuroITV*. Vol. 2006 (2006).
- [20] See-To, Eric WK, Savvas Papagiannidis, and Vincent Cho. "User experience on mobile video appreciation: How to engross users and to enhance their enjoyment in watching mobile video clips." *Technological Forecasting and Social Change* (2012).
- [21] T. Schubert, F. Friedmann, and H. Regenbrecht. "The experience of presence: Factor analytic insights." *Presence: Teleoperators and virtual environments* 10(3), (2001).
- [22] H. O'Brien, and E. G. Toms. "The development and evaluation of a survey to measure user engagement." *Journal of the American Society for Information Science and Technology* 61(1), 50-69 (2010).
- [23] ITU, "Methodology for the subjective assessment of the quality of television pictures." in *Recommendation BT 500-13*, ed: International Telecommunication Union. (2012)

- [24] Schwarz, Heiko, Detlev Marpe, and Thomas Wiegand. "Overview of the scalable video coding extension of the H. 264/AVC standard." *Circuits and Systems for Video Technology*, IEEE Transactions on 17(9), 1103-1120 (2007).
- [25] Gulliver, Stephen R., and Gheorghita Ghinea. "Stars in their eyes: What eye-tracking reveals about multimedia perceptual quality." *Systems, Man and Cybernetics, Part A: Systems and Humans*, IEEE Transactions on 34(4), 472-482 (2004).
- [26] O'Brien, Heather L., and Elaine G. Toms. "What is user engagement? A conceptual framework for defining user engagement with technology." *Journal of the American Society for Information Science and Technology* 59(6), 93 (2008).
- [27] Cortina, Jose M. "What is coefficient alpha? An examination of theory and applications." *Journal of applied psychology* 78(1), (1993).

APPENDIX

The QoVE questionnaire adopted for the proposed study.

<p>Enjoyment (7-point likert scale, 1 = fully disagree, 7 = fully agree)</p>
<p>E1 The video clip was interesting</p> <p>E2 The video clip was fun</p> <p>E3 The video clip was exciting</p> <p>E4 The video clip was enjoyable</p>
<p>Involvement (7-point likert scale, 1 = fully disagree, 7 = fully agree, except for item I1, where 1 = extremely aware, 7 = not aware at all)</p>
<p>I1 How aware were you of the real world surrounding while watching the video clip?</p> <p>I2 When I was watching the video, I lost track of the world around me</p> <p>I3 When I was watching the video, I still paid attention to the world around me*</p> <p>I4 I was completely captivated by the video clip</p>
<p>Satisfaction (7-point likert scale, 1 = fully disagree, 7 = fully agree)</p>
<p>S1 I was satisfied with the experience in watching this video clip</p> <p>S2 I was pleased with the experience in watching this video clip</p> <p>S3 I was contented with the experience in watching this video clip</p> <p>S4 I was delighted with the experience in watching this video clip</p>
<p>Endurability (7-point likert scale, 1 = fully disagree, 7 = fully agree)</p>
<p>D1 Watching this video clip was worthwhile</p> <p>D2 This watching experience was rewarding</p> <p>D3 I want to watch this video clip again</p> <p>D4 I would recommend this video clip to my friends</p>
<p>PVQ (5-point scale [23], with 1 = very annoying, 5 = imperceptible for item PVQ1 and 1 = bad, 5 = excellent for item PVQ2)</p>
<p>PVQ1 Did you perceive any visual impairments in the video?</p> <p>PVQ2 How did you think about the video quality?</p>

*The score of item I3 had been reversed to keep the consistency.