

# Towards the Influence of Audio Quality on Gaming Quality of Experience

Steven Schmidt\*, Saman Zadtootaghaj\*, Shijie Wang\*, Sebastian Möller\*†

\*Quality and Usability Lab, Technische Universität Berlin, Germany, forename.surname@tu-berlin.de

†DFKI Projektbüro Berlin, Germany, sebastian.moeller@dfki.de

**Abstract**—Humans are fascinated by video games for many years, which intrinsically immerse players in their virtual environments. Apart from the challenges offered by steadily new game concepts, aesthetically pleasing environments and characters, stories, and sound effects are highly important for the player experience. Recently, the new concept of cloud gaming, which offers users to play games executed on a cloud server remotely, is becoming increasingly popular. While cloud gaming offers many advantages, additional audio and video streaming poses many technical challenges. To ensure a satisfying Quality of Experience (QoE) of their customers, all stakeholders are interested in finding out which aspects of a gaming experience are of high importance and how resources can be optimally allocated. However, gaming QoE is a multidimensional construct including hedonic and pragmatic aspects and could be strongly influenced by interaction quality, video quality, and audio quality. While the impact of network and video encoding parameters of cloud gaming services was investigated in much detail in recent years, not many studies about the effect of audio quality on gaming experiences are available. Thus, in this paper, the impact of audio quality on gaming experience under different bitrate and packet loss conditions using two popular games is investigated. Therefore, a subjective experiment adhering to the ITU-T Rec. P.809 was conducted. The results show a significant impact of packet loss on audio quality and the overall gaming QoE. However, no significant effect of the bitrate, which was reduced to a minimum of 32 kbps, was revealed. Additionally, the influence of audio quality on gaming QoE was stronger for a game, which contained mainly diegetic effect sounds compared to a game containing few diegetic sounds and affect sound as background music.

**Index Terms**—Gaming, QoE, Audio Quality

## I. INTRODUCTION

The gaming industry has unprecedented managed to motivate users to interact with their services intrinsically. According to the latest report of Newzoo, a leading provider of market intelligence of the global gaming industry, there will be a total of 2.7 billion players across the globe by the end of 2020. The global games market will generate revenues of \$159.3 billion in 2020, an increase of 9.3% compared to the year before [1]. While games like Pong in the 70s were very abstract games, today's games create highly immersive experiences due to their challenges, stories, sound effects, and music. Game music created by professional orchestras such as the *Legend of Zelda* or *Nier* series even fill concert halls and are a fundamental aspect for the success of these games.

978-1-6654-3589-5/21/\$31.00 ©2021 IEEE

While traditionally, games are executed directly on the user's client, e.g., a desktop PC, gaming console, or mobile phone, a new paradigm called *cloud gaming* is gaining increasing interest in the research community and gaming industry. Currently, services such as Google Stadia, XCloud, and GeForce Now are available. According to ITU-T Rec. G.1032, "cloud gaming is characterized by game content delivered from a server to a client as a video stream with game controls sent from the client to the server. The execution of the game logic, rendering of the virtual scene, and video encoding are performed at the server, while the client is responsible for video decoding and capturing of client input" [2].

As this paradigm relies on network performances, the gaming Quality of Experience (QoE) can be strongly influenced by network degradations such as delay and packet loss. Thus, network providers need to ensure fast and reliable connections between the gaming servers and the users' clients. As a network planning tool aiming to predict gaming QoE based on various network and encoding parameters, the recently published opinion model ITU-T G.1072 [3] is available. The recommendation is based on the knowledge summarized in the related ITU-T G.1032 [2], and P.809 [4] recommendations. However, the model so far does not consider the impact of reduced audio quality, even though, as highlighted before, music and sound seem to be of very high importance for a gaming experience. While lots of research is available on the impact of network impairments such as delay on the interaction quality during cloud gaming, it is not well investigated how strong the audio quality contributes to gaming QoE and if there are factors that moderate its impact.

The present paper examines the impact of two common system influencing factors in a cloud gaming service related to audio quality, namely the *bitrate* and *packet loss*. As the game content has proven to be of high relevance when investigating the impact of impairments on gaming QoE, two different games are used in the present research. The cloud gaming setup was used so that the encoding/network impairments only influenced the audio quality, but not the transmission of commands to control the game or the video quality. The paper aims to answer the following research questions:

- Does a reduced bit rate or packet loss have an impact on the user's perceived audio quality?
- Does a reduced audio quality impact the overall gaming QoE and other quality features?

- Is the impact of bitrate or packet loss dependent on the game content?

The remainder of this paper is organized as follows. In Section II an overview of related research is given. Next, in Section III the methodology and test setup for a subjective experiment will be described. Experiment results and a statistical analysis is provided in Section IV. Finally, Section V provides a discussion of the findings and possible future work.

## II. RELATED WORK

In this section, a short overview of research in the audio quality domain in respect to gaming is provided. This information is by no means comprehensive but should allow the reader to understand the decision made in the design of a subjective experiment and a general understanding of the motivation behind the work.

Audio quality is an essential fundament for multimedia services. The audio quality in a telecommunication system may get affected by one of the following processes: recording, encoding, transmission, and post-processing. For cloud gaming service, as the provider has no control over the recording and post-processing phases, the degradation due to encoding and transmission is of interest. Depending on the audio codec and coding parameters' selection, the coding process may significantly degrade the audio quality of a player. Typical audio artifacts that are the result of audio coding could be recognized by multiple distortion type such as quantization noise, binaural unmasking distortions, aliasing artifacts, timbre distortion (birdies), muffled audio (band- limitation), pre-echoes, and so forth [5]–[7]. For further information about the degradation, we refer the reader to [7], chapter 16. One of the most frequently used codecs typically used in gaming streaming platforms is the MPEG-2/4 AAC codec [8], [9] which, similar to most other audio codecs, uses a time-frequency analysis in terms of some transformations to the frequency domain, a so-called modified discrete cosine transform (MDCT). A low-delay version of the AAC codec was proposed, which uses a shorter MDCT transform to reach a low delay of 20ms [7]. Transmissions over IP-based network may degrade the audio quality by introducing errors due to packet loss or concealment artifact (e.g., interruption, frame repetition) in case of unreliable transport via UDP, or stalling and interruptions of the audio stream in case of reliable transport using TCP.

Quality assessment of audio quality is typically done by means of subjective tests. However, service and network providers often utilize prediction models using information extracted from the transmitted signal or network to estimate the audio quality. One of the most successful audio quality models is the standardized audio quality model developed in ITU-T Study Group 12, ITU P.1201 series [10]. The ITU-T P.1201 recommendation series were developed considering different types of codecs, bitrates, packet loss concealment methods, several audio channels, and audio frames to predict the audio quality. We refer the reader to [11] for an overview of existing audio and speech quality models.

The impact of audio quality on gaming experience dimensions, especially the sense of presence, was investigated in some researches. However, most do not consider network or encoding artifacts. In [12]–[14] studies about the impact of audio on the sense of presence are presented. Hedke et al. found a positive effect of spatial audio presentation on the gaming quality of experience [12]. The spatial audio was referred to head-tracked binaural audio presentation based on non-individual head-related transfer functions. While most of the research studies showed an effect of audio quality on gaming quality dimensions such as the sense of presence, to the best of the authors' knowledge, the effect of audio compression and transmission artifact on gaming QoE has not been investigated well in the literature.

When considering the effect of audio degradation on gaming QoE, one should not neglect the difference in the types of sound in a game. Sounds in video games have a wide range of categories, purposes, and triggers. The most widely accepted typologies of game sound is the IEZA model [15] which is a two-dimensional framework. One dimension (diegetic - non-diegetic) makes a distinction between sound emanating from the fictional game world, such as the footsteps of a game character, and sound coming from outside the fictional game world, such as a musical score. The second dimension (setting - activity) makes a distinction on whether the sound is related to the activity in the game or not. By these distinctions, four types of game sound emerge. *Zone* is a diegetic sound related to the setting of the game like the geographical, cultural and/or topological setting of the game world. *Effect* is a diegetic sound related to the activity of the game; which typically is directly synced to a player's activity and game events in the diegetic part of the game environment, such as the sounds of the avatar (i.e., footsteps, breathing), characters dialog, weapons (gunshots, swords), or vehicles (engines, car horns). *Affect* is a non-diegetic sound related to the setting of the game such as orchestral music in an adventure game and horror sound effects in a survival horror game. *Interface* is a non-diegetic sound related to the activity of the game. In many games, Interface contains sounds related to the HUD (Heads-Up Display) such as sounds synced to health and status bars, pop-up menus and the score display.

## III. METHODS

### A. Participants

To answer the research questions, a subjective experiment was conducted. A total number of 26 participants (10 female, 16 male, M=29.3 years, ranging between 20 to 36 years) were tested. Asked about their gaming expertise, 54% of participants described themselves as "intermediate", 27% as "novice" and 19% as "expert" gamers. The majority of participants were foremost PC gamers.

### B. Experimental Design and Manipulations

The experiment followed a within-subject design with three within-subject factors. The first factor, *bitrate* (three levels: 192 kbps, 48 kbps, and 32 kbps), was manipulated by the

encoding process. The second factor, *packet loss* (five levels: 0%, 1%, 4%, 10%, and 20%), was manipulated using a network emulator influencing the network transmitting the audio signal while the bitrate was kept constant at a value of 192 kbps. As a reference condition, a bitrate of 192 kbps and no packet loss is considered in the following. The values for the parameters were selected based on the results of a pre-test, in which participants should solely judge whether they can perceive noticeable differences of the audio quality between each condition (listening-only test). The third factor, the *game* (two levels: Hearthstone, and Counter Strike), was changed by presenting a different game for all seven combinations of the other independent variables. This resulted in a total of 14 different conditions. The stimuli were randomized using a Latin square design to avoid ordering effects.

The dependent variables included the overall gaming QoE, audio quality, input quality, self-judgement of playing performance, as well as seven features of player experience (PX): positive affect, negative affect, tension, competency, challenge, flow, and immersion.

The input quality, also called playability or interaction quality, is measured as the the mean of the quality features responsiveness (e.g., "My inputs were applied smoothly."), controllability (e.g., "I felt that I had control over my interaction with the system."), and immediate feedback (e.g., "I received immediate feedback on my actions."). An overview of the concrete items for the input quality features can be found in [16].

For the assessment of the PX features, the in-game Game Experience Questionnaire (iGEQ) [17] was used. The questionnaire uses two items per feature. In addition, overall gaming QoE and audio quality were assessed using the extended 7-point continuous rating scale (with the labels "extremely bad", "bad", "poor", "fair", "good", "excellent", "ideal"), as proposed in [4]. To avoid confusion by using differently designed scales, we adjusted the discrete 5-point iGEQ scale (with labels "not at all", "slightly", "moderately", "fairly", "extremely") by also applying the extended continuous scale. An example of the scale is shown in Fig. 1.



Fig. 1: Seven-point extended continuous rating scale for iGEQ. This example shows an item to assess challenge. Note that due to the overflow area of the scale, the first label "not at all" was coded as a 2 whereas the label "extremely" was coded as a 6.

### C. Experimental Setup

During the experiment, participants played two different games, namely Hearthstone™(HS) and Counter-Strike: Global Offensive™(CSGO), under different conditions described above. HS is a turn-based card playing game with a constant affect sound as background music creating an

exciting atmosphere. Thus, a reduced audio quality may be permanently noticeable. Additionally, whenever the player interacts with an object, either hovering on a card or dealing a card, there are different effect sounds as feedback. The game also uses a human voice to convey information to the player, such as that a card cannot be played due to a violation of requirements. CSGO is a first-person shooting game with barely any non-diegetic sounds. Even though there is no frequently audible commentator in contrast to HS, CSGO offers many effect sounds due to the players' shootings and footsteps, which can help to identify the positions of opponents. Thus, one could expect a stronger influence on the actual gameplay as the audio effects may improve the performance of players.

To provide impairments on the audio but not on the video or controls of the cloud gaming service, the participants played the game on the server computer (Intel i7 7700K processor, 16GB RAM, and an Nvidia GeForce GTX 1070 graphics card). The games were configured to a 1080p resolution at 60 fps. The audio was transmitted to the client computer (Intel i5 4460 processor, 32GB RAM, and Nvidia GeForce GTX 1060 graphics card) using the Open Broadcaster Software (OBS)<sup>1</sup>. Specifically, the ffmpeg software with AAC audio codec, one of the most commonly used audio codecs, was used. An 24" Asus VG248QE was used as a monitor. Furthermore, the test was conducted in a quiet 83 m<sup>3</sup> acoustically damped room with RT60=0.23 s at 2 kHz. The participants employed the headphones Shure SRH840 (frequency response 5 - 25,000 Hz) with diotic listening to listen to the stimuli. Even though the headphones can be seen as entry level audio equipment, which may influence the perception of audio compared to highly professional hardware, the used equipment can be considered a typical device used by gamers.

To generate different packet loss conditions, the network emulator Clumsy<sup>2</sup> was used on the server computer to control the packet loss rate. Simultaneously, a small delay of 25ms was added to prevent unrealistic error correction available only in a local network. Even though using Clumsy and uniform delay and packet loss might not represent real network characteristics, we consider these simulations sufficient as we solely attempted to create a noticeable difference in audio quality conditions.

To the best of the author's knowledge, the AAC codec is using a signal-adaptive spectrally shaped noise generation for the packet loss concealment [18]. It must be noted that different packet loss concealment are not investigated in this study, which is a limitation when interpreting the concrete quantitative effect of the simulated network impairments. Thus, the focus of this work is the relationships between quality features and gaming QoE.

### D. Experimental Procedure

The experimental procedure followed ITU-T Rec. P.809 [4] and is visualized in Fig. 2.

<sup>1</sup><https://obsproject.com/de>

<sup>2</sup><https://jagt.github.io/clumsy/>



Fig. 2: Structure of experimental procedure.

The experiment started with general questions about demographics and video gaming preferences. Afterward, participants played a training condition of the first game to learn the game’s controls. After finishing the training, the participants played the game under all seven conditions for the selected game. The procedure was then repeated for the remaining game. This block design was used to avoid a time-consuming switching of the games and not confuse the participants. Each stimulus duration was 90 seconds long as recommended by ITU-T Rec. P.809. Participants started a stimulus with a keypress, which executed a script that paused the game by fading the screen to gray after 90 s to avoid an abrupt break potentially influencing immersion-related aspects. After each condition, the participants answered digital questionnaires on a separate computer.

#### IV. RESULTS

##### A. Impact of Bitrate on Audio Quality

To examine the influence of *bitrate* and *game* on subjective ratings of audio quality, a two-way repeated-measures analyses of variance (ANOVA) was computed. The mean values for audio quality are depicted as a bar plot in Fig. 3.

Based on the test statistics, no main effect of bitrate on audio quality was found,  $F(2,50)=0.42$ ,  $p=.66$ ,  $\eta_p^2=.02$ . Also, between the games, no significant main effect was revealed,  $F(1,25)=0.04$ ,  $p=.85$ ,  $\eta_p^2<.01$ . Even though the participants in the pre-test claimed to notice a more narrow audio, the selected range of bitrate did not cause significant differences in the perceived audio quality ratings in the subjective experiment. Likewise, for the overall gaming QoE and the remaining quality features, no main effect of bitrate was found, as visually illustrated for some features in Fig. 3.

These results are surprising considering that also a very low bitrate was used. A discussion about this finding is provided in Section V. Nevertheless, due to the missing impact on the subjective ratings, the analysis will focus on the effect of packet loss.

##### B. Impact of Packet Loss on Gaming QoE, Audio Quality, and Input Quality

To examine the influence of *packet loss* and *game* on subjective ratings of gaming QoE, audio quality, and input quality, multiple two-way repeated-measures ANOVAs were computed, followed by planned contrasts to the reference condition. The ANOVA results for the dependent variables considering packet loss and game as within-subject factors

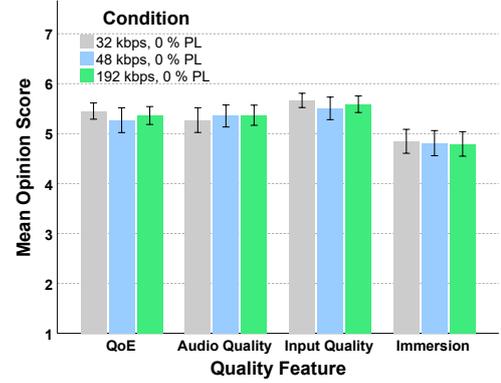


Fig. 3: Bar plots of means and 95 % CI for audio quality averaged over both games influenced by encoding bitrate.

are listed in Table I. The mean values for gaming QoE, audio quality, and input quality are depicted as a bar plot in Fig. 4.

Contrary to bitrate, the ANOVA revealed a significant main effect of packet loss as well as game on *gaming QoE*. Furthermore, an interaction effect of packet loss and game was revealed. Whereas for HS, no significant contrast at 4 % and 10 % packet loss is evident, a significant difference for the gaming QoE ratings was shown for CSGO,  $F(1,25)=9.1$ ,  $p=.006$ ,  $\eta_p^2=.27$  and  $F(1,25)=27.3$ ,  $p<.001$ ,  $\eta_p^2=.52$ , respectively. For HS, significant contrasts started at a packet loss rate of 10 %.

For *audio quality*, the ANOVA yielded comparable effects at the same packet loss rates levels, even though it must be noted that the decrease in the mean values per condition was slightly stronger than for the overall QoE. Furthermore, once again significant differences between the games can be seen.

Regarding the *input quality*, the interaction effect of packet loss and game was much lower than the audio quality but still significant. A significant contrast was reached here at 20 %,  $F(1,25)=10.7$ ,  $p<.003$ ,  $\eta_p^2=.30$ . A significant main effect was shown the first time when comparing the 10 % packet loss condition in contrast to the reference condition,  $F(1,25)=31.4$ ,  $p<.001$ ,  $\eta_p^2=.56$ .

##### C. Impact of Packet Loss on Player Experience

While for the overall gaming QoE, audio quality, and input quality significant effects were revealed for a packet loss rate of 4 % or higher, it is interesting to investigate how a reduced audio quality caused by packet loss would affect the PX features assessed with the iGEQ as well as the self-judgement of a player’s performance. While the corresponding ANOVA test statistics are given in Table I, the mean values in the form of a bar plot are illustrated in Fig. 5.

Apart from the challenge ratings, for all PX features a significant main effect of packet loss was found. However, for *challenge*, a significant main effect of the game was revealed as participants rated HS ( $M=4.11$ ,  $SD=1.07$ ) to be generally more challenging compared to CSGO ( $M=3.62$ ,  $SD=1.09$ ) averaged over all test conditions. This might be explained

TABLE I: Main effects of packet loss and game as well as their interaction effect based on results of two-way repeated measure ANOVA. In cases of violation of sphericity based on Mauchly’s test, degrees of freedom are corrected using Greenhouse-Geisser (for epsilon smaller .75) or Huynh-Feldt correction. The effect size is denoted by  $\eta_p^2$ . Highlighted p-values indicate a significant effect considering an alpha level of .05.

Factor	F	Packet Loss				$\eta_p^2$	F	Game				$\eta_p^2$	Packet Loss x Game				$\eta_p^2$
		$df_1$	$df_2$	p	$\eta_p^2$			$df_1$	$df_2$	p	$\eta_p^2$		F	$df_1$	$df_2$	p	
Gaming QoE	45.25	2.61	65.14	<.001	.64	15.41	1	25	<.001	.38	8.09	4	100	<.001	.24		
Audio Quality	54.37	2.60	64.92	<.001	.69	15.08	1	25	<.001	.38	3.58	4	100	.01	.13		
Responsiveness	15.78	2.51	62.67	<.001	.39	3.95	1	25	.06	.14	2.32	4	100	.06	.08		
Controllability	15.75	2.05	51.23	<.001	.39	0.88	1	25	.36	.03	3.19	2.71	67.66	.03	.11		
Immediate Feedback	22.23	2.55	63.65	<.001	.47	2.70	1	25	.11	.10	2.44	2.46	61.38	.08	.09		
Input Quality	2.38	2.39	59.72	<.001	.45	2.80	1	25	.11	.10	3.11	2.81	7.22	.03	.11		
Immersion	23.01	2.08	52.00	<.001	.48	4.10	1	25	.05	.14	3.69	4	100	.01	.13		
Competency	11.07	4	100	<.001	.31	0.40	1	25	.53	.02	3.45	2.97	74.35	.02	.12		
Negative Affect	12.22	2.64	66.07	<.001	.33	5.77	1	25	.02	.19	2.09	4	100	.09	.08		
Flow	14.42	2.52	63.06	<.001	.37	8.42	1	25	.01	.25	2.12	4	100	.08	.08		
Tension	16.02	2.59	64.74	<.001	.39	0.45	1	25	.51	.02	2.81	4	100	.03	.10		
Positive Affect	19.54	2.64	65.90	<.001	.44	1.91	1	25	.18	.07	2.34	3.55	88.87	.07	.09		
Challenge	0.82	4	100	.51	.03	18.05	1	25	<.001	.42	0.23	4	100	.92	.01		
Playing Performance	4.91	4	100	<.001	.16	0.15	1	25	.71	.01	1.59	4	100	.18	.06		

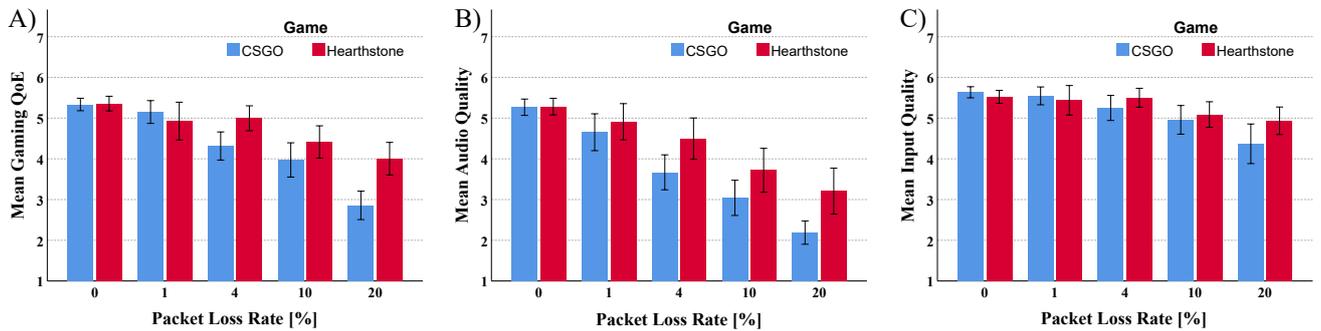


Fig. 4: Bar plots of means and 95 % confidence interval for overall gaming QoE (A), audio quality (B), and input quality (C) for both games influenced by packet loss.

by the report in the post-game questionnaire stating that the learnability and feedback about the performance by the game were lower for CSGO.

For *immersion*, *tension*, and *competency*, a significant interaction effect of packet loss and game was found. This means that for one game, the impact of packet loss was different than for the other game. In particular, for CSGO, the immersion ratings are already significantly reduced at a packet loss rate of 4% compared to the reference condition, whereas for HS, a significant effect was found starting at 10%. Additionally, for HS, there was no difference between the 1% loss condition with neither 4% nor 10%, but for CSGO each of these comparisons revealed significant differences,  $p < .01$ . For tension, the interaction effect was caused by the 20% loss condition, as this subjective ratings were significantly higher in CSGO compared to all other conditions,  $p < .01$ , whereas for HS only the reference condition was significantly lower rated,  $p < .05$ . For competency, significant differences emerged for CSGO already at 4% loss, whereas packet loss had no effect for any conditions during HS.

Interestingly, on the contrary to the expectations, only a tendency of an interaction effect for the *playing performance*

feature was shown. For CSGO, only when comparing the reference ( $M=4.96$ ,  $SD=0.91$ ) with the 20% loss condition ( $M=4.16$ ,  $SD=1.26$ ),  $p=02$ , a significant effect on performance was revealed.

Regarding the main effect of packet loss on *negative affect*, *flow*, and *positive affect* ratings, it can be reported that the contrast comparison to the reference condition was first significant for the 4% loss condition for all three quality features.

To investigate the relationships between the quality features, a correlation analysis was performed across both games. The results show significant relationships between audio quality and subjective measures for *overall gaming QoE* ( $r=.80$ ,  $p < .01$ ), *input quality* ( $r=.42$ ,  $p < .01$ ), *immersion* ( $r=.62$ ,  $p < .01$ ), *competency* ( $r=.20$ ,  $p < .01$ ), *negative affect* ( $r=-.33$ ,  $p < .01$ ), *flow* ( $r=.42$ ,  $p < .01$ ), *tension* ( $r=-.49$ ,  $p < .01$ ), *positive affect* ( $r=.45$ ,  $p < .01$ ), as well as with *playing performance* ( $r=.21$ ,  $p < .01$ ). Whereas the relationship between audio quality and immersion was the highest among the PX features, the relationship to playing performance is rather low.

## V. DISCUSSION AND CONCLUSION

In the presented research, we investigated the impact of two common system influence factors, the encoding bitrate

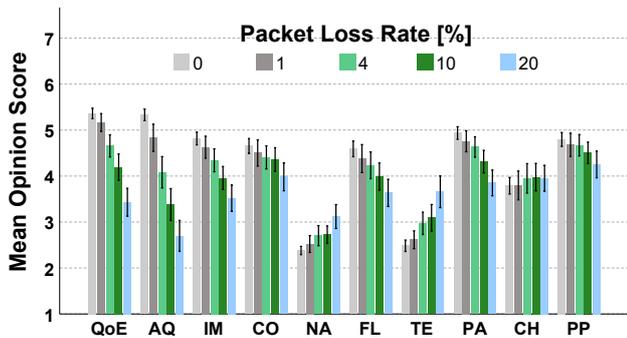


Fig. 5: Bar plots of means and 95 % CI for overall gaming QoE, audio quality (AQ), immersion (IM), competency (CO), negative affect (NA), flow (FL), tension (TE), positive affect (PA), challenge (CH), and playing performance (PP) influenced by packet loss averaged for both games.

and packet loss, on the audio quality and gaming QoE in the context of a cloud gaming service.

The results showed no effect of bitrate even at the lowest bitrate of 32 kbps. This was initially a surprising finding. However, one must consider that audio quality was only one of many assessed quality features. It might be that during the highly engaging gaming activity, the audio degradations due to a reduced bitrate did not get a high attention of participants. Even though the results were revealed for both games, more research is required to confirm these findings. It would be also interesting, whether similar results would be found when inviting core gamers as test participants.

When comparing subjective ratings of audio quality and a variety of PX features, starting at a packet loss rate of 4 %, significant differences were found for the majority of features. Interestingly, the impact of packet loss was higher for the shooting game CSGO, which frequently uses effect sounds due to shootings and footsteps of the players. Here, the input quality was more negatively impacted as players might have noticed issues in the game’s auditive feedback when performing time-critical interactions. Furthermore, contrary to HS, the effect sounds typically contribute to an improved localization of enemies in CSGO and, thus, provide a benefit for more successful gameplay.

Regarding the modeling of gaming QoE, such as in the recently published ITU-T Rec. G.1072, for which audio quality was not considered, this research’s findings indicate that audio quality in the tested range of parameters can also significantly affect gaming QoE and underlying features. However, as in the presented study very high packet loss rates were used, more research is required to investigate the quantitative effect for a higher resolution of the parameters. Here, especially an investigation of packet loss on the audio and video simultaneously is of interest since it is likely that at the same packet loss rate, the reduced video quality is more impactful on the overall gaming QoE compared to reduced audio quality. Furthermore, when attempting to include audio quality in gaming QoE model, details about different codecs and packet loss concealment

methods must be taken into account. The same applies to a more realistic simulation of the network conditions in respect to their distribution. In the future, we plan to address these limitations to improve the generalizability of the findings.

#### ACKNOWLEDGMENT

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 871793.

#### REFERENCES

- [1] T. Wijman, “The World’s 2.7 Billion Gamers Will Spend \$159.3 Billion on Games in 2020; The Market Will Surpass \$200 Billion by 2023,” Available at: <https://newzoo.com/insights/articles/newzoo-games-market-numbers-revenues-and-audience-2020-2023/>, 2020.
- [2] ITU-T Recommendation G.1032, “Influence Factors on Gaming Quality of Experience,” International Telecommunication Union, Geneva, 2017.
- [3] ITU-T Recommendation G.1072, “Opinion Model Predicting Gaming QoE for Cloud Gaming Services,” International Telecommunication Union, Geneva, 2020.
- [4] ITU-T Recommendation P.809, “Subjective Evaluation Methods for Gaming Quality,” International Telecommunication Union, Geneva, 2018.
- [5] M. Erne, “Perceptual Audio Coders” What to Listen for,” in *Audio Engineering Society Convention 111*. Audio Engineering Society, 2001.
- [6] C.-M. Liu, H.-W. Hsu, and W.-C. Lee, “Compression Artifacts in Perceptual Audio Coding,” *IEEE transactions on audio, speech, and language processing*, vol. 16, no. 4, pp. 681–695, 2008.
- [7] S. Möller and A. Raake, *Quality of Experience: Advanced Concepts, Applications and Methods*. Springer, 2014.
- [8] ITU-T Rec. H.262, “Generic coding of moving pictures and associated audio,” 1994.
- [9] M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, and M. Dietz, “ISO/IEC MPEG-2 Advanced Audio Coding,” *Journal of the Audio Engineering Society*, vol. 45, no. 10, pp. 789–814, 1997.
- [10] M.-N. Garcia, P. List, S. Argyropoulos, D. Lindgren, M. Pettersson, B. Feiten, J. Gustafsson, and A. Raake, “Parametric Model for Audiovisual Quality Assessment in IPTV: ITU-T Rec. P. 1201.2,” in *2013 IEEE 15th International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 2013, pp. 482–487.
- [11] A. W. Rix, J. G. Beerends, D.-S. Kim, P. Kroon, and O. Ghitza, “Objective Assessment of Speech and Audio Quality—technology and Applications,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 6, pp. 1890–1901, 2006.
- [12] T. Hedke, J. Ahrens, J. Beyer, and S. Möller, “Impact of Spatial Audio Presentation on the Quality of Experience of Computer Games,” *Proc. of Jahrestagung fr Akustik (DAGA’17)*, p. 4, 2017.
- [13] P. Skalski, R. Whitbred, and P. Lindmark, “Image vs. Sound: A Comparison of Formal Feature Effects on Presence, Video Game Enjoyment, and Player Performance,” in *12th Annual International Workshop on Presence*, 2009.
- [14] J. Lessiter, J. Freeman, and J. Davidoff, “Really hear? The Effects of Audio Quality on Presence,” in *Proceedings of the Fourth Annual International Workshop on Presence*, 2001, pp. 288–324.
- [15] R. van Tol and S. Huiberts, “IEZA: A Framework for Game Audio,” *Retrieved January*, vol. 21, p. 2013, 2008.
- [16] S. Schmidt, “Assessing the Quality of Experience of Cloud Gaming Services,” Ph.D. dissertation, TU Berlin, 2021.
- [17] W. IJsselstein, Y. De Kort, and K. Poels, *The Game Experience Questionnaire*. Eindhoven: Technische Universiteit Eindhoven, 2013.
- [18] H. Autti and J. Biström, “Mobile Audio—from MP3 to AAC and further,” *Helsinki University of Technology*, 2004.