User Quality of Experience Comparison Between Skype, Microsoft Teams and Zoom Videoconferencing Tools

Marko Matulin[1](✉), Štefica Mrvelj[1], Borna Abramović[1], Tomislav Šoštarić[1], and Marko Čejvan[2]

[1] Faculty of Transport and Traffic Sciences, University of Zagreb, Vukeliceva 4, 10000 Zagreb, Croatia
{mmatulin,smrvelj,babramovic}@fpz.hr
[2] Ericsson Nikola Tesla d.d, Krapinska 45, 10000 Zagreb, Croatia
marko.cejvan@ericsson.com

Abstract. In 2020, we investigated the quality of online meetings to find out whether various video and audio degradations can affect end-users' perception and experience. We collected a total of 542 questionnaires that were suitable for analysis. This paper uses the results to compare the Quality of Experience for three popular videoconferencing applications: Skype, Microsoft Teams, and Zoom. Due to the COVID-19 (Coronavirus disease 2019) and anti-epidemic measures, survey participants were all faced with using these applications daily, whether for work, attending lectures, or keeping in touch with friends and family. They rated their frustration level for specific quality degradations (e.g., blocking of the image or echo in the audio). The paper shows how the three applications compare to each other when different network performance degradation situations occur and how they affect the end user's perception.

Keywords: Videoconference · Application · Online meetings · Telecommuting · Quality of experience · Survey

1 Introduction

The unique circumstances created by the coronavirus disease forced humanity to accept different living and working ways that were rather unthinkable in the past. Probably the most significant change took place in our work environments when lockdown measures led to various videoconferencing applications coming to the fore. It is safe to say that there is no industry that does not use the above tools (often referred to as telecommuting) to at least some extent, and certain sectors, such as education, rely on them. According to [1, 2], usage of these tools began to rise significantly in 2010, and since then it has more than doubled. It is expected that by 2030, about 50% of the working population will participate in videoconferencing.

T. Šoštarić—Graduated with a masters' degree.

D. Perakovic and L. Knapcikova (Eds.): FABULOUS 2021, LNICST 382, pp. 299–307, 2021.
https://doi.org/10.1007/978-3-030-78459-1_22

Telecommuting has great potential to reduce costs, mainly for transportation to and from work/school/faculty (up to 30% according to [3]), as well as business operating costs. Zoom, a videoconferencing platform, currently holds the largest share in this market. At the end of 2019, Zoom was averaging 10 million users/month. In April 2020, the number peaked at 300 million daily meeting participants. Dozens of other videoconferencing applications on the market make it a competitive market with high growth expected in the coming years. Therefore, it was interesting to compare how the three most popular applications (according to our data) behave when various degradations in video and audio quality occur during online meeting sessions.

As the name suggests, video is the essential component of videoconferencing service. Nevertheless, there are crucial differences when comparing this service with, e.g., video streaming. During the conferencing session, video is not stored before playback; it is encoded and streamed live from a conference to a group of participants located at different network sites, often with limited upload/download capabilities. The limitations can affect the quality of the video transmission and thus the Quality of Experience (QoE) for all conference peers. Therefore, it is important to understand how users react in situations where network performance declines, leading to various video and audio quality degradations on the end-user side.

To this end, researchers focus their energy on understanding both objective and subjective aspects of videoconferencing services. This is in line with the QoE concept of service evaluation, which puts end-users into the focus and investigates how quantifiable (objective) and qualitative (subjective) parameters impact their perception. Network performances that can meet a specific application's demands are still an integral part of successful service provision, but they are not the only ones; users' subjective opinions come to the fore in the QoE paradigm of service evaluation.

The available research on the objective characteristics of the service includes a) video coding techniques for videoconferencing [4, 5], b) the development of adaptive video streaming methods [4, 6, 7], or c) the analysis of mobile audiovisual telemeetings with multiple participants [8–10]. However, our focus in this paper is on subjective aspects of user QoE, i.e., their perception of different quality degradations when using popular videoconferencing applications. Note that this follows up on our earlier work from [11], where we presented descriptive statistics of our survey results.

Belmudez in [12] presented the QoE model for videoconferencing services, which uses objective parameters to derive subjective user opinions about the service. Similar work but for a video streaming service is done in [13]. Another attempt to discover the interplay between network performance and user QoE for videoconferencing is done in [14], where Perceived Video Quality (PVQ) was used to evaluate the user experience. The results on how different coding strategies in video communication scenes affect subjective perception were published in [15] by Li et al. The authors investigated the applicability of varying encoding strategies to different video classes.

In [16], the authors investigate how varying network conditions affect four subjects' ability to collaborate and build a Lego model using a videoconferencing platform. Rao et al. [17] developed different scenarios to perturb network performance on the sender side. They quantified the perceived video quality at the receiver side but without performing subjective tests with actual test subjects.

Before moving on to the study design and presenting the results, we are obliged to highlight a crucial aspect of our research that distinguishes it from other research efforts reviewed here. We could not measure network performance during online meetings or the audiovisual quality of our survey participants' incoming videos as we approached them through social networks and personal channels. The survey was conducted outside of laboratory conditions without actual physical contact with respondents. Therefore, we did not focus on examining the quality of a particular session or inbound video stream. Instead, we revealed respondents' opinions about the quality of their typical videoconferencing session when using different applications.

As mentioned earlier, we are continuing our previous work [11], where we revealed and discussed how frequent various quality degradations are and how they affect end-users (e.g., different audio and video artifacts and meeting disconnections). In this paper, we are focused on discovering if the user perception is in interplay with the videoconferencing platform type, i.e., how different platforms perform in the network performance deprivation scenarios (from an end-user perspective).

This paper is organized as follows. Section 2 describes the questionnaire used in the survey and general statistics about the sample. The results obtained and discussions can be found in Sect. 3. Section 4 concludes and offers new research paths.

2 Subjective Study

A detailed description of the survey, the questionnaire used to collect the data, and the survey participants can be found in [11]; therefore, we aim to provide a more summary report in this paper.

We used the LimeSurvey platform and designed a questionnaire with 31 questions that took about 15 min to complete. In addition to demographic data, we collected responses on a) respondents' network environment from which they typically access meetings (device and network types, experiences with, e.g., long buffering times when viewing the video stream and other audiovisual quality degradations), b) typical meeting scenarios (meeting purposes, what videoconferencing applications they typically use, etc.) and meeting roles respondents were involved in, and c) respondents' opinions on meeting quality. The survey lasted three weeks and was anonymous.

After excluding uncompleted questionnaires and those where respondents indicated that they had not attended meetings, we finally obtained 542 questionnaires that were suitable for analysis. For this paper, we further filtered the data as follows: a) only respondents who attended more than five meetings in the last month were included (measured from the time they participated in the survey); b) respondents ranked over 100 applications based on how frequently they used them. We filtered the three top-ranked applications and performed data analysis based on those responses.

The results showed that the top three most frequently ranked applications were Zoom, Microsoft Teams[1], and Skype[2] (ranked as the number one choice by 104, 96, and 49 respondents, respectively).

[1] Hereinafter refered as Teams.

[2] Note that the distinction was made between Skype and Skype for Business.

As seen from Table 1, the sample is sufficiently diversified (given the gender and age group distributions). For each application studied here, over 60% of respondents participated in more than ten meetings in a month, making them experienced users. We also learned that the responders' devices are usually connected to a home network (DSL or WiFi) during the meetings. Based on the responses related to how long it takes for a typical video streaming playback to start, how often it re-buffers, and how

Table 1. Sample characteristics.

Sample data	Skype	Teams	Zoom
A number of responses:	49	96	104
Gender:			
Female	63.3%	38.5%	59.6%
Male	36.7%	61.5%	40.4%
Age group:			
19–30	6.1%	24%	7.7%
31–40	36.7%	16.7%	30.8%
41–50	40.8%	35.4%	37.5%
51–60	12.2%	16.7%	19.2%
61–70	4.1%	6.3%	3.8%
71 or older	0%	1%	1%
The number of attended meetings:			
Between 6 and 10	38.8%	27.1%	22.1%
Between 11 and 20	28.6%	26.0%	37.5%
Between 21 and 40	20.4%	27.1%	27.9%
41 or more	12.2%	19.8%	12.5%
Most frequently used device for meetings:			
Laptop	73.5%	80.2%	67.3%
Smartphone	8.2%	9.4%	15.4%
Desktop computer	16.3%	8.3%	14.4%
Tablet	2%	2.1%	2.9%
Common meeting purposes:			
Work-related	42.9%	67.7%	50.0%
To attend or give lectures	34.7%	29.2%	38.5%
To see friends and relatives	12.2%	3.1%	6.7%
Other	10.2%	0%	4.8%
Typical meeting role:			
Meeting leader and presenter	42.9%	35.4%	45.2%
Participant, sometimes a presenter	38.8%	37.5%	30.8%
Participant	12.2%	17.7%	14.4%
Presenter	2%	4.2%	6.7%
Guest	0%	1%	1.9%
Meeting leader but others present	4.1%	4.2%	1%

often the resolution changes, we conclude that the respondents' network was sufficiently capacitated in most cases for the videoconferencing service.

3 Results and Discussion

This section presents the results on three figures that follow the same structure and have six subplots, two for every application we analyze. Figure 1 and 2 show respondents' frustration levels for audio and video degradations, respectively, while Fig. 3 compares these frustrations to the most adversely perceived quality distortion – a meeting disconnections. The cumulative distribution function (CDF) was used to present the results. Five-point discrete scales were used to quantify the distortion frequency during one typical conference and describe the respondents' frustration levels.

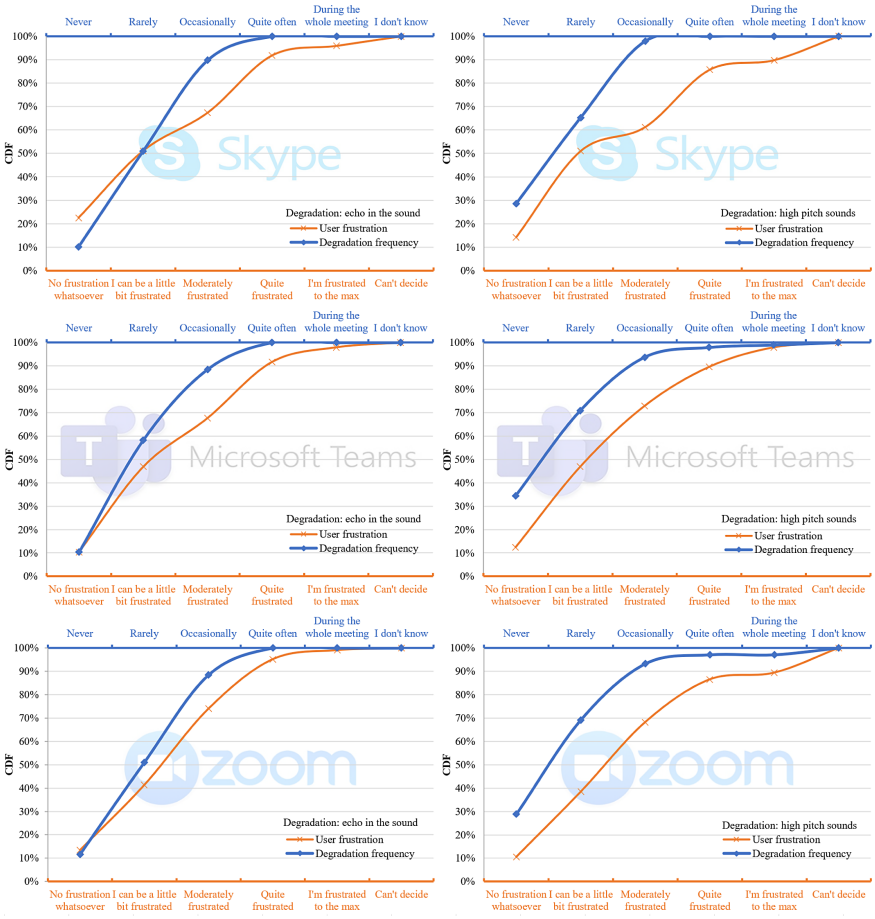

Fig. 1. Responders' frustration for a specific type of audio degradation and its' frequency.

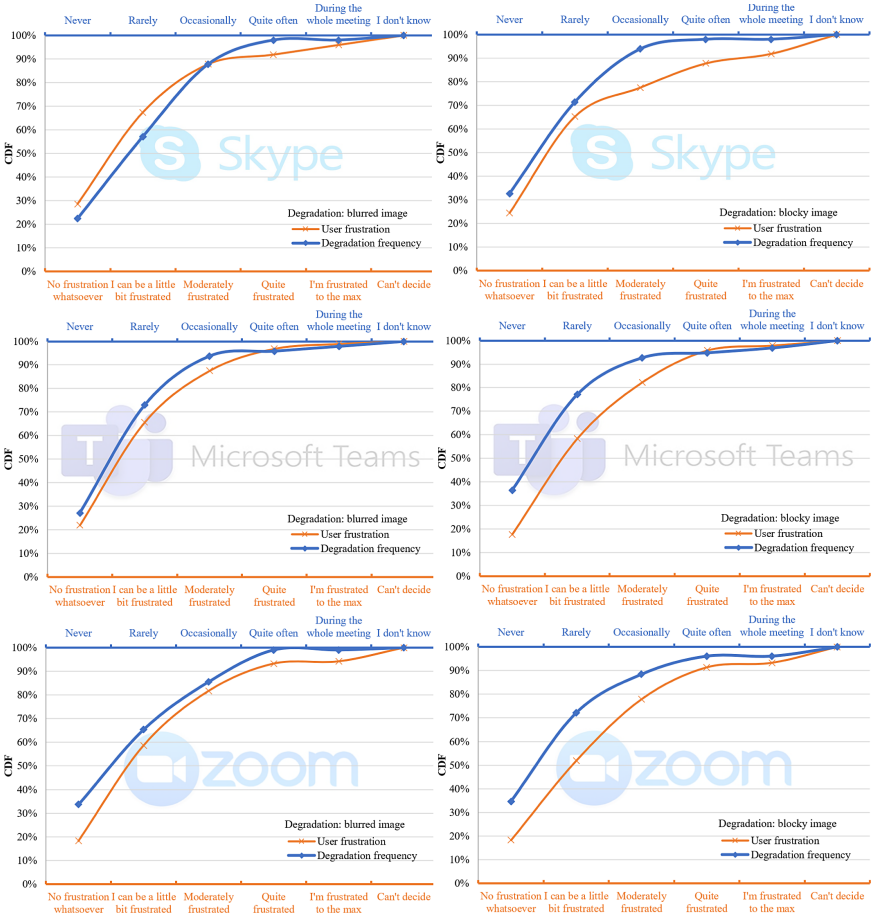

Fig. 2. Responders' frustration for a specific type of image degradation and its' frequency.

Note that in [11], the results showed that users value audio more than image quality. The same conclusion can be drawn here, for this filtered sample, if we compare the results shown in Fig. 1 and Fig. 2. For example, the value of CDF for *moderate* levels of frustration on Skype is 67.4% (for echo in the sound) and 61.2% (for high pitch sound) compared to 87.8% (for image blurriness) and 77.6% (for image blockiness). The same trends are seen for the other two applications.

The relationship between the frequency of degradation and user frustration is somewhat different for Skype. We can see that a higher level of frustration was recorded for this particular population. It is arguably most evident for *high pitch sound* (Fig. 1), where 61.2% of respondents indicated that this degradation was *not frustrating*, or could be a *little frustrating*, or was *moderately frustrating*. Conversely, the CDF values for Teams and Zoom, for the same *moderate* frustration level, are higher, i.e., there are fewer respondents who indicated higher levels of frustration (*quite frustrated* and *maximally frustrated*).

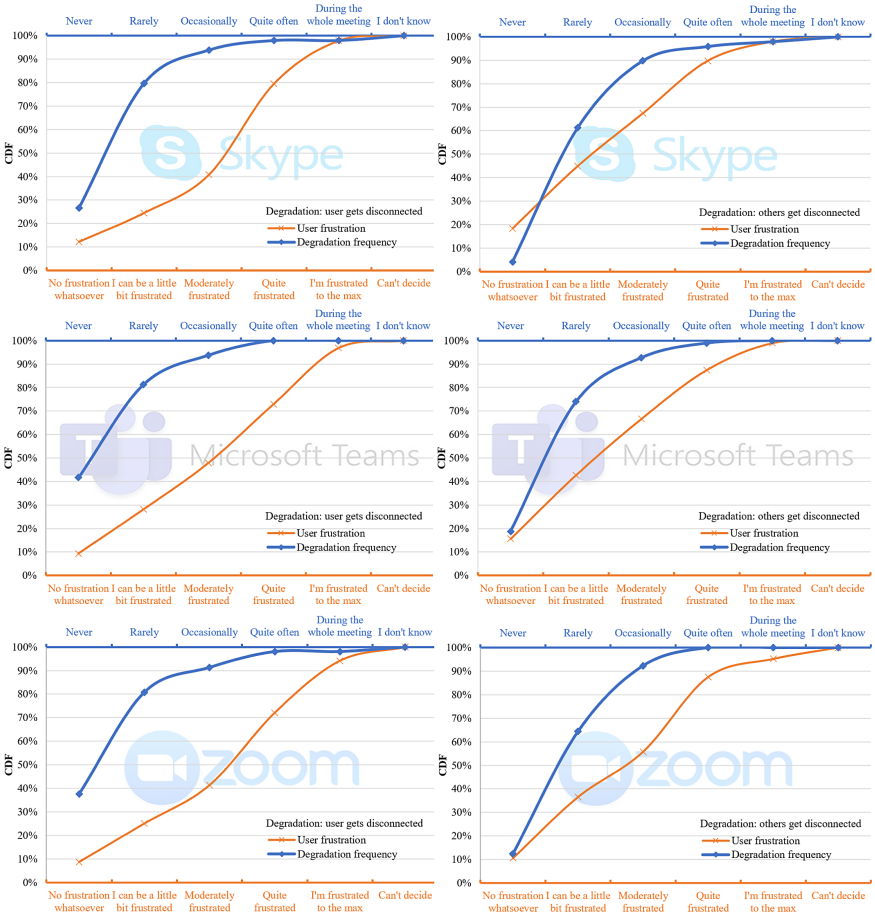

Fig. 3. Responders' frustration for a meeting disconnections and their frequency.

Respondents' attitudes towards the meeting disconnections is the last set of results to be presented here. Figure 3 shows the values of CDF for two types of degradations, namely when a user (a respondent) is disconnected from a conference (three subplots on the left) and when other conference peers are disconnected (the subplots on the right).

It is clear that the frustration values visualized in Fig. 3 are significantly higher than those we have previously reported for audio or video degradations. Of the three applications studied, Teams users were the least frustrated by this degradation, which is probably because the frequency of this degradation is slightly lower for Teams than Skype and Zoom. On the other hand, this result can be interpreted to mean that Zoom users are more accustomed to holding high-quality meetings, so these types of meeting interruptions are more likely to be perceived negatively. It can also be seen that in all three applications, more frustration is elicited when the responders were disconnected than when their conference peers were disconnected.

To test the differences between multiple sample proportions, we used the Chi-square test. We recorded a statistically significant difference between CDF values for *occasional* degradation frequency only for image related degradations. In contrast, the CDF values were statistically significantly different for *moderate* levels of frustrations for all degradations and between all three applications we have analyzed.

4 Conclusion

This paper tried to reveal correlations between different levels of user dissatisfaction and frequency of specific quality degradations for specific videoconferencing applications. We detected the three most frequently used applications and conducted the data analysis only for that filtered sample based on the survey results.

The results showed that disconnections during the meetings are most adversely perceived quality degradation, especially for Skype and Zoom users. Nearly 60% of these users reported higher levels of frustration with this type of degradation in cases where they were disconnected. It shows that even though various applications may provide similar functionalities, they can create different user experiences, which tells us that this research effort is worthwhile.

In the future, we will focus on performing more sophisticated statistical analyses that can uncover a deeper understanding of user perception and experience for videoconferencing. We also plan to explore the interplay between various objective and subjective factors of user QoE for this type of service.

Conflicts of Interest. The authors declare that there is no conflict of interest. The results were not influenced in any way by any organization mentioned in the manuscript.

References

1. GrandViewResearch. https://www.grandviewresearch.com/industry-analysis/video-conferencing-market. Accessed 05 June 2020
2. Gminsights. https://www.gminsights.com/industry-analysis/video-conferencing-market. Accessed 05 June 2020
3. Skillscouter. https://skillscouter.com/video-conferencing-statistics. Accessed 06 June 2020
4. Bakar, G., Kirmizioglu, R.A., Tekalp, A.M.: Motion-based rate adaptation in WebRTC videoconferencing using scalable video coding. IEEE Trans. Multimedia **21**(2), 429–441 (2019). https://doi.org/10.1109/TMM.2018.2856629
5. Boyce, J.M., Ye, Y., Chen, J.L., Ramasubramonian, A.K.: Overview of SHVC: scalable extensions of the high efficiency video coding standard. IEEE Trans. Circ. Syst. Video Technol. **26**(1), 20–34 (2016). https://doi.org/10.1109/TCSVT.2015.2461951
6. Hu, H., Zhu, X.Q., Wang, Y., Pan, R., Zhu, J., Bonomi, F.: Proxy-based multi-stream scalable video adaptation over wireless networks using subjective quality and rate models. IEEE Trans. Multimedia **15**(7), 1638–1652 (2013). https://doi.org/10.1109/TMM.2013.2266092
7. Hosking, B., Agrafiotis, D., Bull, D., Easton, N.: An adaptive resolution rate control method for intra coding in HEVC. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, China, pp. 1486–1490. IEEE (2016)

8. Li, CL., Xiong, HK., Zou, JN., Wu, DPO.: Joint dynamic rate control and transmission scheduling for scalable video multirate multicast over wireless networks. IEEE Trans. Multimedia **20**(2), 361–378 (2018). https://doi.org/10.1109/TMM.2017.2745709
9. Al Hasrouty, C., Lamali, M.L., Autefage, V., Olariu, C., Magoni, D., Murphy, J.: Adaptive multicast streaming for videoconferences on software-defined networks. Comput. Commun. **132**, 42–55 (2018). https://doi.org/10.1016/j.comcom.2018.09.009
10. Vučić, D., Skorin-Kapov, L.: QoE assessment of mobile multiparty audiovisual telemeetings. IEEE Access **8**, 107669–107684 (2020). https://doi.org/10.1109/ACCESS.2020.3000467
11. Matulin, M., Mrvelj, Š., Abramović, B.: How frustrated are you? User perception about different videoconference quality degradations. In: 5th EAI International Conference on Management of Manufacturing Systems (EAI MMS 2020), EAI, Online (2020)
12. Belmudez, B.: Audiovisual Quality Assessment and Prediction for Videotelephony. Springer, Berlin (2015). https://doi.org/10.1007/978-3-319-14166-4
13. Matulin, M., Mrvelj, Š.: Modelling user quality of experience from objective and subjective data sets using fuzzy logic. Multimedia Syst. **24**(6), 645–667 (2018). https://doi.org/10.1007/s00530-018-0590-0
14. ur Rehman Laghari, K., Issa, O., Speranza, F., Falk, T.H.: Quality-of-experience perception for video streaming services: preliminary subjective and objective results. In: Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference, Los Angeles, USA, pp. 1–9 (2012)
15. Li, Y., Zhou, Z.: Subjective video quality assessment and the analysis of coding strategies in video communication scene. In: Proceedings of The 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Chengdu, China, pp. 52–55 (2020). https://doi.org/10.1109/CISP-BMEI51763.2020.9263575
16. Schmitt, M., Redi, J., Bulterman, D., Cesar, P.S.: Towards individual QoE for multiparty videoconferencing. IEEE Trans. Multimedia **20**(7), 1781–1795 (2018). https://doi.org/10.1109/TMM.2017.2777466
17. Rao, N., Maleki, A., Chen, F., Chen, W., Zhang, C., Kaur, K., Haque, A.: Analysis of the effect of QoS on video conferencing QoE. In: 15th International Wireless Communications & Mobile Computing Conference (IWCMC), Tangier, Morocco, pp. 1267–1272 (2019). https://doi.org/10.1109/IWCMC.2019.8766591