electronics

MDPI

Article

Modeling the Level of User Frustration for the Impaired Telemeeting Service Using User Frustration Susceptibility Index (UFSI)

Štefica Mrvelj and Marko Matulin *

Faculty of Transport and Traffic Sciences, University of Zagreb, 10000 Zagreb, Croatia; smrvelj@fpz.unizg.hr
* Correspondence: mmatulin@fpz.unizg.hr

Abstract: Modern users are accustomed to always-accessible networks ready to serve all of their communication, entertainment, information, and other needs, at the touch of their devices. Spoiled with choices provided on the competitive markets, the risk of customer churn makes network and service providers sensitive to user Quality of Experience (QoE). Services that enable people to work and industries to function in these pandemic times, such as the telemeeting service, are becoming ever more critical, not just for the end-users but also for the providers. Nevertheless, the heterogeneity of end-users network environments and the uniqueness of the service (bidirectional video and audio transmissions and interactivity between the meeting peers) imposes specific QoE requirements. Hence, this paper focuses on understanding how different service quality degradations affect user perception and frustration with such impaired service. The impact of eight quality degradations was analyzed. Based on the conducted user study, we used the multiple regression analysis and developed three models capable of predicting user Level of Frustration (LoF) for the specific degradations that we have analyzed. The models work with the User Frustration Susceptibility Index (UFSI), which categorizes users into groups based on their tendency to become frustrated with the impaired service.

Keywords: quality of experience; telemeeting; videoconference; audiovisual; quality; impairments; degradations; user studies; frustration susceptibility; modeling

Citation: Mrvelj, Š.; Matulin, M. Modeling the Level of User Frustration for the Impaired Telemeeting Service Using User Frustration Susceptibility Index (UFSI). *Electronics* **2021**, *10*, 2202. https://doi.org/10.3390/electronics10182202

Academic Editor: Paul Mitchell

Received: 7 August 2021
Accepted: 7 September 2021
Published: 9 September 2021

1. Introduction

Because of the unique circumstances imposed by the pandemic COVID-19, societies worldwide were forced to rethink and reshape their everyday work and living habits. In adapting to the new conditions, many industries were pushed to use online-based tools that otherwise would not have been used, at least not to the extent they were due to the pandemic. The videoconferencing service is a perfect example of such a tool; many sectors use it frequently, while some sectors, such as education, rely on it under these new-normal conditions.

Scrolling through recent reports on the size of the videoconferencing market (e.g., [1,2]) yields interesting information. For instance, the market size has more than doubled since 2010. About half of the working population is expected to participate in multiparty telemeetings in this decade. One of the main incentives for using this technology is the potential cost reduction, especially commuting costs, the reduction in operating costs for businesses, more efficient time management, more flexible work schedules, etc. There are dozens of video conferencing platforms vying for market position, with Zoom being the current market leader. Therefore, it is safe to say that this is a competitive industry that is expected to grow at a high rate in the coming years (measured in billions of dollars). Predictively, this growth is already accelerated by the pandemic and the uncertainty surrounding its end.

Before turning to the survey of related work on telemeeting quality evaluation and user perceptions, it is beneficial to review the concept of Quality of Experience (QoE) and its definition. The QoE idea emerged at the beginning of the new millennium when various multimedia applications became widely available over the Internet. Different network requirements were defined for each type of application. These are called QoS (Quality of Service) requirements and are typically used for network design and management. For example, Internet telephony quality is considered acceptable if packet latency is kept below 150 ms. Network operators and service providers were thus given a set of quantifiable, measurable goals per specific application type that should ensure satisfied customers. However, it was found that users are sometimes not fully satisfied with services even when QoS requirements have been met.

Several authors attempted to show why the relationship between measurable QoS requirements (i.e., network performance) and user perception is not as straightforward as expected. It was found that many qualitative, subjective factors that are difficult to measure significantly influence users' opinions about the service. These factors include the social context of service usage, internal user state, and their prior experiences, quality of content, application design, and numerous others [3]. Therefore, a new concept (QoE) was introduced, providing a holistic approach to telecommunication service quality analysis.

The starting point to know and understand QoE today is found in [4], from where we provide the following QoE definition: "Quality of Experience is the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state." As can be seen from the definition, the user is the focus of the quality analysis. Network performances that can meet the QoS requirements of a given application are still an integral part of a successful service delivery but not the only one. In the QoE paradigm of service evaluation, users come to the forefront. Surveying their subjective opinions (in person or online) becomes an essential method in the researcher's toolbox.

In light of the frequently cited QoE definition from [4], we focus our study on learning the degree to which users are annoyed when various quality disturbances occur during their telemeeting sessions. Specifically, we analyzed the impact of eight types of service degradations on users. These are the echo in audio, audio with high pitch, audio containing noise, video blurriness, video blockiness, video interruptions, user disconnections, and disconnections of other meeting peers. Although there have been past research efforts analyzing user perceptions for some quality impairments, to the best of our knowledge, this is the first attempt to disclose user perceptions and internal state related to the above group of service impairments. Moreover, based on our data collection and analysis, we developed regression models capable of predicting the levels of user frustrations for different service impairments, taking into account users' susceptibility to frustrations and several qualitative factors.

The following sections are: Section 2 reviews the current research on QoE in videoconferencing and discusses our motivation for conducting this experiment; Section 3 explains the study design and sample characteristics; Section 4 presents and discusses the results of the user study, including the modeling, while Section 5 concludes.

2. Literature Review and Motivation

Video is an essential component of a video conferencing service. Nevertheless, there are crucial differences when comparing this service with, for example, video streaming. During video conferencing, the video is not stored before playback; instead, it is encoded and streamed live from one network location to:

- another user, i.e., two participants in total are in a meeting which makes it a telemeeting session, or,
- more than one user, i.e., >2 participants are in a meeting, making it a multiparty telemeeting session [5].

The participants are located at different network locations, often limited by bandwidth. The limitations can affect the video transmission quality and the QoE for the participants in that heterogeneous environment. Therefore, it is crucial to understand how users react when the network performance degrades, leading to various degradations in audiovisual quality at the end-user side. To this end, researchers focus on understanding both the objective and subjective aspects of telemeeting service quality. This is in line with the QoE concept of service evaluation, which focuses on the end-user and examines how quantifiable (objective) and qualitative (subjective) parameters affect their perception.

Since the success of this service depends on video quality, considerable efforts have been made to derive optimized methods and techniques for video distribution. For example, in [6], video coding techniques for videoconferencing are proposed using rate adaptation through motion-based spatial and temporal resolution selection. The authors claim that their solution allows each peer to send a video to different peers with different terminal types and network rates using a single encoder.

In addition to [6], adaptive video streaming methods were also the focus of [7], where Hu et al. develop a scalable video adaptation over wireless networks to address the heterogeneity of both video streams and the underlying wireless link capacities of different users. The network heterogeneity problem was also analyzed in [8] by Al Hasrouty et al., who propose two algorithms for establishing and maintaining multicast sessions in a software-defined network. Li et al. in [9] developed a scheme to maximize the overall video reception quality among all destinations during the multicast period. The scheme significantly improved the network multicast throughput and showed the flexibility to dynamic network changes.

In the work reviewed thus far [6–9], researchers focused on the network side of a telemeeting quality, i.e., they invested efforts in optimizing video transmission in different network ecosystems. However, as mentioned in the introduction, this is only one part of the QoE evaluation. To understand how a service performs in a QoE approach, one needs to survey end-user opinions and perceptions. Performing only QoS-to-MOS (Mean Opinion Score) mapping, for instance, is not sufficient for the service operators since the QoE is a much more complex issue [10]. Hence, this second part of the literature review on user QoE for telemeetings includes [11], where several aspects related to the quality perception of video calls were addressed, namely, the quality estimation in an interactive context, the audiovisual quality integration of single audio and video modalities, and the temporal pooling of short sample-based quality scores to account for the perceptual quality impact of time-varying degradations. Additionally, the QoE model is presented, which uses the packet loss rate to derive user QoE level for this type of service.

Laghari et al. in [12] used the Perceived Video Quality (PVQ) metric to evaluate user perception for different videoconference distortion scenarios. The scenarios included various packet loss rates (0, 0.5, 1, 3, 7, and 15%), packet reordering (0, 1, 5, 10, 20, and 30%), and coding bit rates (100, 400, 600, and 800 Kbps). They showed an exponential relationship between PVQ and packet loss and between PVQ and packet reordering, while a logarithmic relationship was established between PVQ and video bit rate. Similar to [12], Rao et al. in [13] introduced jitter, packet loss, uplink throttling, and latency into their distortion scenarios and used DMOS (Differential Mean Opinion Score) to evaluate incoming video signal from the end-user perspective, thus quantifying the QoS-QoE relationship. They evaluated 144 conferencing sessions under predefined network scenarios and found that the bandwidth is directly proportional to the perceived quality of the video. Users preferred a steady bandwidth over irregularly increasing bandwidth and could forgive increasing jitter and packet loss if the bandwidth stayed stable and high.

An example of using network-related parameters to predict the appearance of audiovisual disturbances during multiparty telemeetings can be found in [14], where a machine learning technique was used for WebRTC performance estimation. In a series of papers published after four years of research, Vučić et al. analyzed the impact of different smartphone configurations [15], video resolutions and bandwidth [16], and packet loss [17] on user QoE for telemeetings in the mobile environment, while investigating the unexpected quality disturbances during WebRTC sessions in [18]. In [19], they found that providing consistent lower objective video quality is better than switching between higher and lower qualities since participants start to perceive impairments. This is in line with the findings presented by Rao et al. in [13], where users highly valued a steady bandwidth.

In [20,21], experiments were conducted to determine how different network conditions affect four test subjects' ability to collaborate and build a block model over a videoconferencing platform. The authors focused on disclosing the subjects' interactions and the level of understanding in a real-life scenario about the task at hand, thus, broadening the QoE context for this type of service. It was found that lower bit rates slow the interaction between the participants who shifted their focus from the video to the audio channel. Additionally, the authors discover how users become more forgiving about quality degradations once a system has enabled them to interact. They concluded by stating that if the objective is to accurately estimate the QoE of participants, knowing the system factors alone does not suffice.

De Moor et al. in [22] and Husić et al. in [23] found that audio quality has the most decisive impact on user perception, compared with other impact factors (e.g., image quality or QoS). Silva et al. in [24] measured user annoyance for image blockiness and blurriness and found that the blockiness had the stronger impact on users between the two. In a recent study [25], Øie et al. investigated a patient's and a doctor's QoE during video consultations. They found that the patient's age, medical condition, expectations, attitudes, prior experiences, and technical skills and competence of both patient and doctor affect their perception. Similar to the experiments conducted in [20,21], the research highlights the importance of studying the qualitative factors of a telemeeting service as per a specific service use case.

In our previous work, we reported our survey of user QoE for videoconferencing. Specifically, we discussed the survey results using only descriptive statistics [26] and showed how users perceive the service if different conferencing platforms were used, namely, Microsoft Teams, Skype, and Zoom [27]. This paper deepens that knowledge by investigating possible relationships between user frustration for eight types of quality distortions (audio and video related impairments and meeting disconnections). Apart from the degradations, the impact of other qualitative factors (such as meeting role, meeting purpose, and others) on user perception was analyzed. Based on the literature review, we believe this is the first time that the impact of such a broad distortion spectrum and the interplay between them on the perception of several hundreds of users are analyzed.

Moreover, we used multiple regression analysis for modeling user frustration levels for different audiovisual impairments and meeting disconnections. The developed models operate with several predictor variables, such as degradation frequency, user opinion about the importance of audio/video quality during the sessions, User Frustration Susceptibility Index (UFSI), and others. We developed UFSI to categorize users into groups based on their tendency to become frustrated with the impaired service. The UFSI also acknowledges recently proposed QoE research approaches for other services, such as [10] for web browsing and video streaming, [28] for video streaming, and [29] for gaming video streaming. In these studies, traditional metrics, such as MOS, are enriched with others derived from complex relations between the QoE influential factors and taking user diversity into account.

3. Study Design and Sample Characteristics

This section explains the concept of our survey and describes the online questionnaire used for data collection. It concludes by discussing the sample characteristics after it was cleared from outliers and inexperienced telemeeting users.

3.1. Concept of the Survey

This survey differs from others because our user perception study did not focus on a single telemeeting session, nor did we attempt to evaluate, for example, the quality of a particular incoming video stream in different network scenarios (as in [11–13]). Due to the pandemic and associated locked-down measures, we could not conduct face-to-face tests and interviews and collect users' opinions about specific session quality. We reached out to survey participants remotely and asked them to provide us with data about their typical telemeeting scenario and quality. After participating in multiple telemeetings over one month, we learned more about respondents' perceptions as they experienced audiovisual disruptions and session interruptions. That is important because we do not attempt to associate a session-specific QoS metric with user QoE. Instead, we report user opinions of service quality after participating in numerous conferencing sessions.

3.2. The Questionnaire Used in the Survey

The questionnaire was hosted on the LimeSurvey platform for four weeks. It contained 31 questions and took about 20 min to complete. We used our previous user studies [30] and personal telemeeting experiences and developed the questionnaire. The questions are grouped into five categories described below.

1. With the first group of questions, we learned whether the respondent had participated in telemeetings last month (measured from the time of participation in our survey). If the answer was Yes, the survey continued, and we asked how many of them and when the last meeting was. If the answer was No, the survey was closed.
2. General demographic data were collected with this question category (respondents' gender, age group, education status, and employment status).
3. With the third set of questions, we learned from where the respondents usually connect to the meetings (home or work or both) and what devices they often connect with (desktop computer, laptop, smartphone, or tablet). In addition, we were interested in finding out what type of network the devices are connected to during the meetings (WiFi, DSL, mobile network, fiber-to-the-home, cable, or satellite network).
4. Here, we asked questions to identify the respondents' most common: (a) meeting purpose (work related or socializing with friends and family), (b) used application (we provided a list of over 100 different items to choose from), and (c) meeting role they are in (host, presenter, participant, or guest).
5. The last category included questions about participants' views on essential parameters for a good quality telemeeting (audio and video quality, screen sharing, application interface, understanding how the role system works, etc.). In addition, we learned what quality degradations usually occur when they are in the meeting and how often (various audiovisual degradations or meeting interruptions) and how frustrating it is for them to experience these degradations.

Different types of questions were used in each category: single- and multiple-choice questions, questions asking respondents to rank, for example, specific aspects of the service in order of importance (audio and image quality, interface design of the application used for conferencing, etc.), and questions asking users to rate their level of frustration per specific quality degradation on a five-point scale with the following ratings: No frustration whatsoever, I can be a little bit frustrated, Moderately frustrated, Quite frustrated, and I'm frustrated to the max. The participants also evaluated how often specific degradations appear during their sessions. That frequency was also rated on a five-point scale with the following linguistic meanings: Never, Rarely, Occasionally, Quite often, and during the whole meeting.

When we felt it was necessary, we briefly explained specific quality degradation and how it can manifest itself in a user application, helping survey participants identify and evaluate how that audiovisual degradation affects them.

3.3. Sample Characteristics

Of 542 completed questionnaires, we excluded those in which participants indicated that they: (a) had been involved in fewer than six telemeetings, (b) did not know if the audiovisual degradation described ever occurred in their sessions, (c) did not know if they experienced any level of frustration during the meetings, and (d) did not appreciate certain aspects of the service (e.g., audio or video quality) but then reported a higher level of frustration when the degradation affecting that aspect of the service occurred (e.g., echo in the audio or video blockiness, respectively). We implemented these exclusion criteria to keep only the answers from more experienced users who could recognize different degradations, express their opinion about them, and give consistent answers. Thus, the analysis continued on 322 questionnaires (see Table 1 for sample population details).

Table 1. General information about the survey participants and their meetings.

Characteristic	Description	Distribution
Gender	women men	48% 52%
Age group	between 19 and 30 between 31 and 40 between 41 and 50 between 51 and 60 between 60 and 70 over 70 years of age	14.86% 30.03% 31.58% 17.65% 5.57% 0.31%
Number of attended telemeetings in a monthly period	between 6 and 10 sessions between 11 and 20 sessions between 21 and 40 sessions over 40 sessions	29.42% 31.58% 25.39% 13.62%
Most frequently used device for the meetings	laptop desktop computer smartphone tablet	69.97% 15.48% 12.07% 2.48%
Common meeting purposes	work related to attend or give lectures to see friends and relatives other	50.77% 38.39% 6.19% 4.64%

This paper aims to analyze the session-specific quality degradations and their effect on telemeeting peers. As expected, not all survey participants noticed all types of degradations that we have studied; hence, in Table 2, we present the number of questionnaires included in the analysis of specific quality degradation. Table 2 also shows that the most commonly noticed degradation was an echo in the audio (N = 287); it is also worthwhile noting that 99 participants reported that they had experienced all types of degradations during their meetings in a monthly period.

Table 2. The number of survey participants that did not notice the specific quality degradation.

Types of Quality Degradations	Did Not Notice the Degradation	Remaining (N)
Echo in the audio	35	287
High-pitch audio	79	243
Audio contains noise	61	261
Video is blurred	82	240
Video is blocky	110	212
Video is interrupted	88	234
I get disconnected	104	218
Others get disconnected	38	284

To determine a possible difference in a sample structure (between the participants who did not notice a particular degradation—sample A, and those who experienced it—sample B), we conducted a hypothesis test for comparing two sample proportions. To this end, a 5% level of significance ($\alpha = 0.05$) was chosen. Figure 1 depicts the sample A and B structure (top and bottom row, respectively) for the echo in the audio degradation (as seen from Table 2, sample A and B size equals 35 and 287, respectively). The figure shows that both samples are similar, and we observed the same similarity for other degradations.

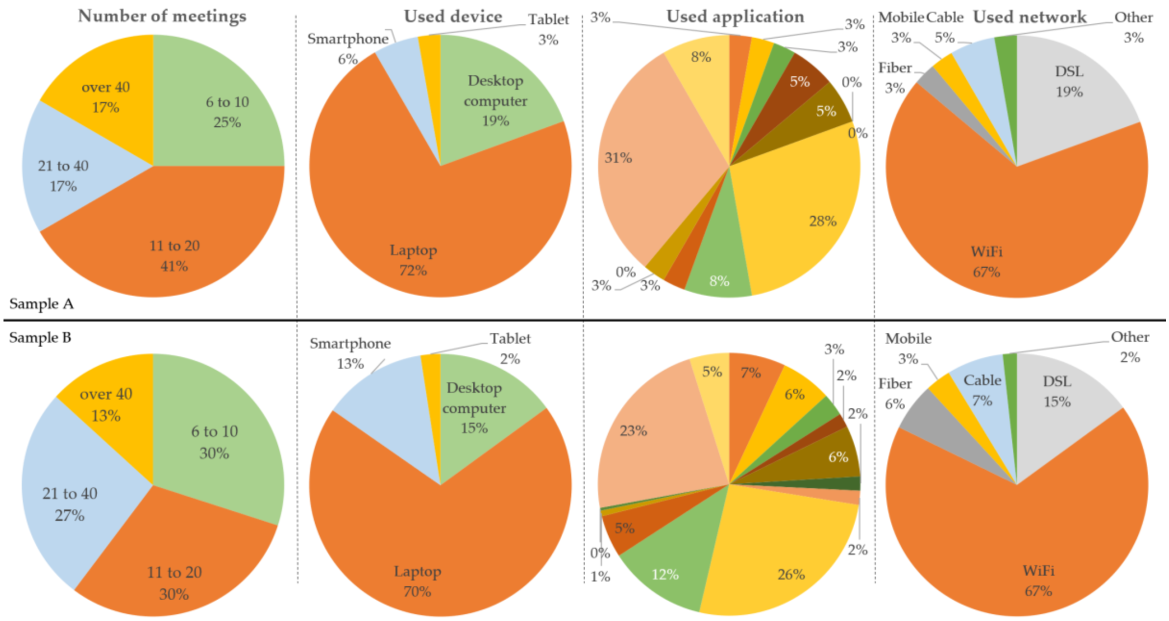

Figure 1. Sample A (top row) and sample B (bottom row) structures were based on the number of attended meetings, devices, applications, and type of network used in telemeetings. For the sake of clarity, in the third column (used application), only the percentages are shown.

A hypothesis test to determine whether the difference between the two proportions is significant showed the following.

- Echo in the audio/video is blurred/video is blocky/video is interrupted. All differences in sample proportions are in the interval of accepting the null hypothesis ($H_0: p_1 - p_2 = 0$), so we cannot reject the null hypothesis and conclude that, for instance, sample A contains a higher percentage of survey participants who attended a certain number of meetings or used a specific type of application than the percentage of participants in sample B.
- High-pitch audio/others get disconnected. We cannot accept the null hypothesis only for the share of participants who used the Microsoft Teams application (a marginally higher share of these participants is in sample A).
- The audio contains noise. We cannot accept the null hypothesis only for the percentage of participants who used the Jitsi application (a slightly higher percentage of them is in sample B).
- I get disconnected. The null hypothesis cannot be accepted only for those participants who attended 11–20 meetings (sample A contains a slightly higher share of them) and those who attended more than 41 meetings (their share is marginally higher in sample B).

4. Results and Model Development

The following chapters present the user study results and the conclusions which they led to. The developed models are also revealed and verified within the scope of this section.

4.1. Definition of Training and Testing Sets

The degradations were grouped into three groups: (1) audio degradations (containing echo in the audio, high-pitch audio, and audio contains noise), (2) video degradations (containing video is blurred, video is blocky, and video is interrupted), and (3) meeting disconnections, i.e., session interruptions (with two degradation types included: a survey participant was disconnected or other telemeeting peers were disconnected). We decided to group the degradations when we discovered that the survey participants affected by one type of service impairment are more susceptible to others from the same group (discussion about the participants' susceptibility to degradations follow in Section 4.3. onward).

Before defining multiple correlation coefficients and regression models per specific quality degradation group, we divided the data set into two subsets: training set (T_r) and test set (T_e). The participants who noticed all three degradations from the audio group of degradations were used for T_r (206 of them) for the degradations of that group, while others were in the T_e. The same was conducted for the video degradations group, where 145 participants experienced all three degradations from that group. Since only three participants reported that they did not notice disconnection of other meeting peers, but they have been disconnected, T_e for the "I get disconnected" degradation type was defined differently (40 participants were selected randomly, and we added 3 participants who did not notice disconnections of others. Details of the T_r and T_e per specific degradation are shown in Table 3.

Table 3. The size of the training (T_r) and testing (T_e) sets per each degradation.

Types of Quality Degradation	T_r	T_e	Total
Echo in the audio	206	81	287
High-pitch audio		37	243
Audio contains noise		55	261
Video is blurred	145	95	240
Video is blocky		67	212
Video is interrupted		89	234
I get disconnected	175	43	218
Others get disconnected	215	69	284

4.2. Multiple Regression Analysis

An analysis was conducted to see how the survey participants' level of frustration for different degradation types (Table 3) was impacted by the equipment they were using, the meeting purpose, the application they were using, etc. For the sake of clarity, we now assign predictor variables to the questions asked in the survey. As discussed in Section 3.2, the participants provided answers about: how frequent specific degradation appeared in their telemeetings (X_f); what was their most common telemeeting purpose (X_p); the roles they assumed in the meetings (X_r); what equipment (X_e) and application did they use (X_{app}), on what type of network (X_n); how important is the audio (X_a), webcam video stream (X_v), and share screen quality (X_s); and how many meetings did they attend (X_m) in a monthly period.

The above questions were posed as predictor variables in determining the association with the Level of Frustration (LoF) set as a response variable depending on the type of degradation, i.e., degradation group (the three groups described in the previous chapter). For different degradation types, the structure of the multiple regression models is shown below: Equations (1)–(3) for audio-LoF_a, video-LoF_v and the disconnection group of degradations-LoF_d, respectively).

$$LoF_a = \beta_0 + \beta_1 \cdot X_p + \beta_2 \cdot X_r + \beta_3 \cdot X_a + \beta_4 \cdot X_m + \beta_5 \cdot X_e + \beta_6 \cdot X_{app} + \beta_7 \cdot X_n + \beta_8 \cdot X_f \tag{1}$$

$$LoF_v = \beta_0 + \beta_1 \cdot X_p + \beta_2 \cdot X_r + \beta_3 \cdot X_v + \beta_4 \cdot X_s + \beta_5 \cdot X_m + \beta_6 \cdot X_e + \beta_7 \cdot X_{app} + \beta_8 \cdot X_n + \beta_9 \cdot X_f \tag{2}$$

$$LoF_d = \beta_0 + \beta_1 \cdot X_p + \beta_2 \cdot X_r + \beta_3 \cdot X_m + \beta_4 \cdot X_e + \beta_5 \cdot X_{app} + \beta_6 \cdot X_n + \beta_7 \cdot X_f \tag{3}$$

While analyzing the data, scatterplots were used to observe relationships between all predictor variables and the response variable. The obtained plots depicted unclear direction, form, and strength of the relationships between the variables. We calculated the correlation coefficients and found that the correlation between each analyzed variable and the level of frustration was small. Apart from the linear correlation, we tested the curvilinear relationship between the variables (exponential, logarithmic, polynomial, and power) and found no significant difference in the strength of the relationships. Hence, we continued using linear regression.

We used matrix techniques and multiple regression analysis in Excel to calculate the linear regression coefficients, their standard errors, multiple correlation coefficients, and additional regression statistics. Multiple correlation coefficients (Multiple R) and multiple coefficients of determination (R Square) for different types of degradation are shown in Tables 4–6. As can be seen from the data shown in the tables, the correlation between these predictor variables and the LoF can be assessed as weak or moderate, depending on the literature source used for interpretation (Figure 2). From Figure 2, we can see that the multiple correlation coefficients are approximately equal, and somewhat higher correlation coefficients were obtained for the two types of video distortions, namely video blurriness and blockiness.

Table 4. The summary output for the audio group of degradations.

Regression Statistics	Echo in the Audio	High-Pitch Audio	Audio Contains Noise
Multiple R	0.40939	0.37694	0.38989
R Square	0.16760	0.14208	0.15202
Standard Error	1.01696	1.00019	0.97972
Observations	206	206	206

Table 5. The summary output for the video group of degradations.

Regression Statistics	Video is Blurred	Video is Blocky	Video is Interrupted
Multiple R	0.44981	0.47098	0.37427
R Square	0.20233	0.22182	0.14008
Standard Error	0.91005	0.94015	1.06349
Observations	145	145	145

Table 6. The summary output for the disconnection group of degradations.

Regression Statistics	I Get Disconnected	Others Get Disconnected
Multiple R	0.34070	0.34286
R Square	0.11608	0.11755
Standard Error	1.14196	1.1308
Observations	175	215

To assess the significance of a given $R^2_{Y.1,\ldots,J}$ (where Y is a dependent variable and J represents the number of independent variables), we computed an F ratio. All calculated values of F are significant at all the usual alpha levels ($\alpha = 0.05, 0.10,$ and 0.01). Therefore, we can reject the null hypothesis ($H_0 : R^2_{Y.1,\ldots,J} = 0$, all $p < 0.05, 0.10,$ and 0.01). Computed F values can be found in Table 7.

Due to the p-value $< 0.05/0.10/0.01 = \alpha$, we conclude that the regression model is a significantly good fit; i.e., there is only a 0.001–0.823% possibility of obtaining a correlation this high (0.34–0.47) assuming that the null hypothesis is true. We also see that R squares are mostly lower than 0.2 (i.e., only 20% of the variance in the level of frustration is explained by the model), and the standard errors of the estimate are in the [0.91, 1.14] interval.

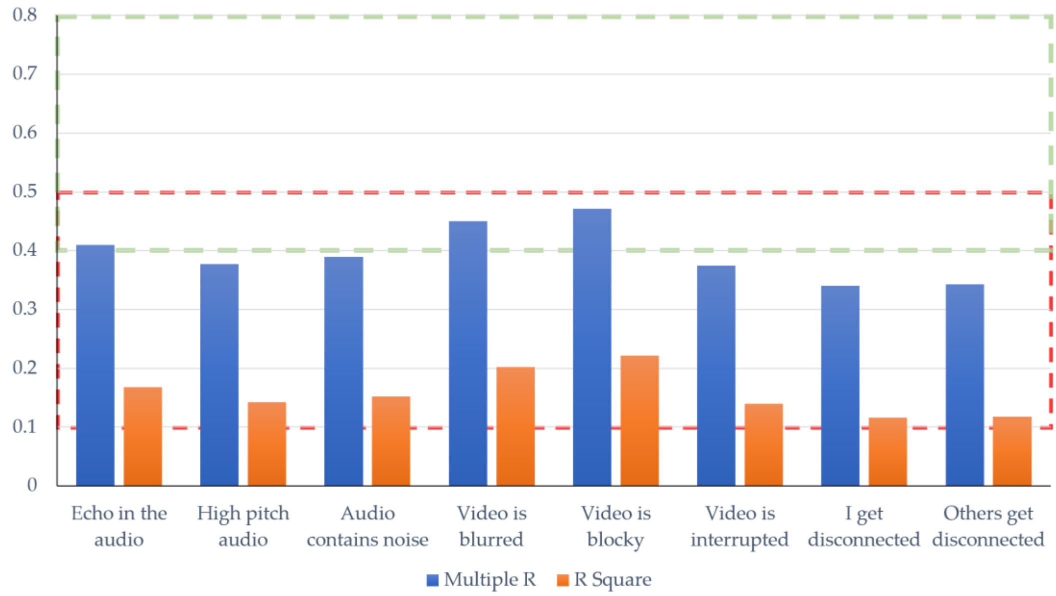

Figure 2. Multiple correlation coefficients (multiple R) and multiple coefficients of determination (R square). The green and red dashed rectangles represent moderate and weak correlations intervals, respectively. The overlapping between the rectangles indicates how different sources provide different interpretations of the intervals.

Table 7. ANOVA table for all types of degradations.

Degradation Type		df	SS	MS	F	Significance F
Echo in the audio	Regression	8	41.02151	5.12769	4.95804	0.00001
	Residual	197	203.74062	1.03422		
	Total	205	244.76214			
High-pitch audio	Regression	8	32.63833	4.07979	4.07823	0.00016
	Residual	197	197.07526	1.00038		
	Total	205	229.71359			
Audio contains noise	Regression	8	33.89840	4.23730	4.41451	0.00006
	Residual	197	189.09189	0.95986		
	Total	205	222.99029			
Video is blurred	Regression	9	28.35939	3.15104	3.80472	0.00026
	Residual	135	111.80613	0.82819		
	Total	144	140.16552			
Video is blocky	Regression	9	34.01408	3.77934	4.27585	0.00007
	Residual	135	119.32385	0.88388		
	Total	144	153.33793			
Video is interrupted	Regression	8	24.48780	3.06097	2.71961	0.00823
	Residual	136	153.07082	1.12552		
	Total	144	177.55862			
I get disconnected	Regression	7	35.26090	5.12769	3.93934	0.00046
	Residual	207	264.69259	1.03422		
	Total	214	299.95349			
Others get disconnected	Regression	7	28.59870	4.08553	3.13292	0.00391
	Residual	167	217.77844	1.30406		
	Total	174	246.37714			

Since correlation coefficients are classified as weak or moderate, regression models give an insufficiently accurate estimation of the level of user frustration by a particular type of degradation. Perhaps, this is best illustrated in Figure 3 that shows only one kind of quality degradation from each degradation group (audio, video, and disconnection). The modeled frustration levels, computed by Equations (1)–(3) and depicted with orange dots, somewhat follow the degradation frequency (gray dashed line). However, we see that the blue line, indicating the participants' frustration levels, does not follow the degradation frequency (gray dashed line), meaning the model provides a poor fit. This suggests a variable(s) whose values were not measured in our survey exist(s), impacting the results. We assume that the reason for such a large discrepancy between the actual and modeled frustration levels is a person's predisposition for being frustrated. The same conclusions can be drawn for the other degradations analyzed.

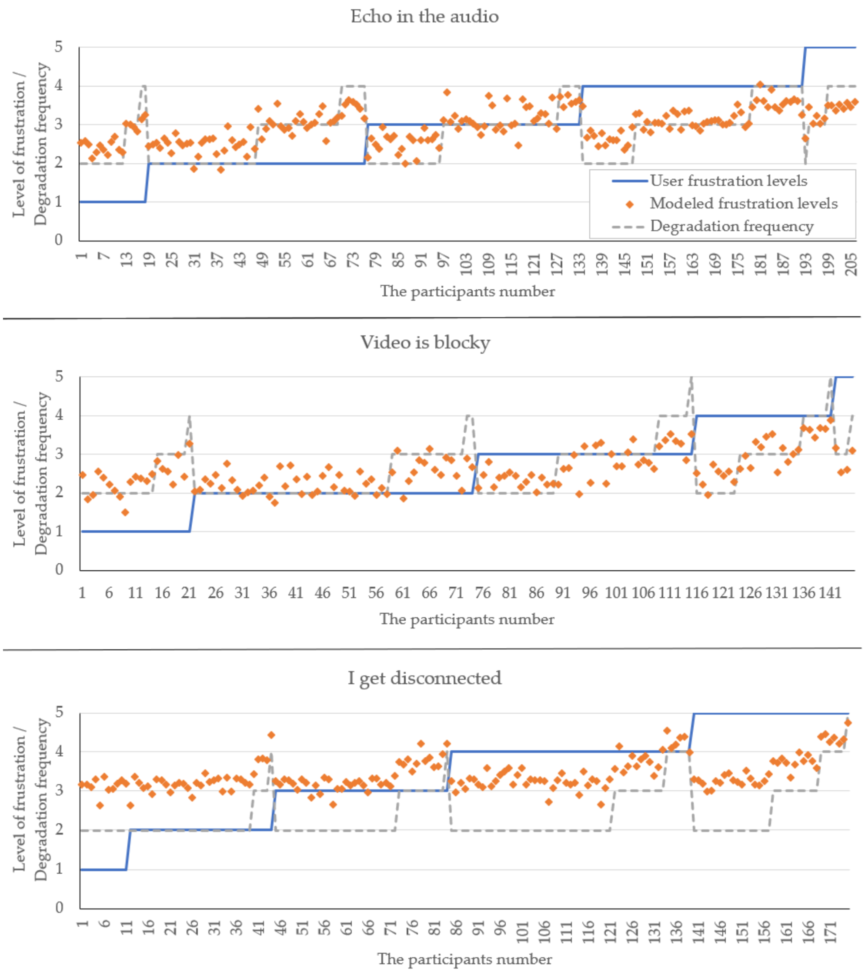

Figure 3. Participants' frustration levels, degradation frequency, and modeled frustration levels for three quality degradation types.

Figure 3 also shows that the highest level of frustration (a response I'm frustrated to the max in a survey), corresponding to a rating of 5 in the charts, is reported only by a few participants for echo in the audio and video is blocky quality degradations. Their number is higher for the third degradation depicted in Figure 3. For that specific degradation, a significant share of participants expressed the maximum level of frustration for all degradation frequency cases. Specifically, the percentage of maximally frustrated users equaled 15% when the degradation frequency was reported as 2 (meaning rarely), 29% when the frequency was 3 (meaning occasionally), 43% for the frequency of 4 (representing quite often), and 100% for the highest frequency of 5 (meaning during the whole meeting).

As seen from all three subplots of Figure 3, most participants reported that the degradations were rarely or occasionally happening, corresponding to the values of 2 and 3 of the *y*-axis of the charts. A small share of participants believes that the frequency was quite often (rating of 4 on the charts), and even fewer participants experienced the degradations during the whole meeting (rating of 5 on the charts).

It is important to note that the assessment of the degradation frequency was entirely a subjective opinion of an individual survey participant. As discussed in Section 3.1, we are not trying to model human perception for specific session-related QoS metrics, nor were we able to measure the participants' network conditions and/or audiovisual quality during their online meeting sessions. This implies there is a level of ambiguity when discussing what rarely, occasionally, etc., means for a particular survey participant.

Hence, in the following chapter, we examine the correlation between frustration levels per individual quality degradation. Namely, we try to confirm our assumption that participants who were frustrated with one type of degradation were more likely to be frustrated with other types.

4.3. Correlation Coefficients between the Frustration Levels

In analyzing the interplay between each type of degradation and the associated level of participants' frustration, we found that their occurrence frequency does affect users. Still, the effect is not as strong as anticipated. This prompted us to investigate other factors that might interfere with their perception. To this end, we decided to examine whether one type of quality degradation could also make participants more susceptible to the other types and more annoyed. Table 8 shows the values of the Pearson's correlation coefficients between the two variables, where the variables are the degree of frustration caused by a particular type of degradation $r_{frustration}$ (the values written in black colored text). The values that are written in grey colored text are correlation coefficients between the degradation frequency per specific degradation type ($r_{frequency}$).

Table 8. The correlation between the levels of frustration and degradation frequency for a pair of quality degradations.

Types of Quality Degradations	Correlations	Types of Quality Degradations							
		Echo in the Audio	High-Pitch Audio	Audio Contains Noise	Video Is Blurred	Video is Blocky	Video Is Interrupted	I Get Disconnected	Others Get Disconnected
Echo in the audio	Corr. coeff. Sig (2-tailed)		0.809 *** 0.000	0.722 *** 0.000	0.554 ** 0.000	0.600 ** 0.000	0.596 ** 0.000	0.483 ** 0.000	0.538 ** 0.000
High-pitch audio	Corr. coeff. Sig (2-tailed)	0.569 ** 0.000		0.740 *** 0.000	0.558 ** 0.000	0.507 ** 0.000	0.512 ** 0.000	0.451 ** 0.000	0.507 ** 0.000
Audio contains noise	Corr. coeff. Sig (2-tailed)	0.441 ** 0.000	0.562 ** 0.000		0.654 ** 0.000	0.638 ** 0.000	0.533 ** 0.000	0.523 ** 0.000	0.509 ** 0.000
Video is blurred	Corr. coeff. Sig (2-tailed)	0.371 * 0.000	0.444 ** 0.000	0.451 ** 0.000		0.848 *** 0.000	0.607 ** 0.000	0.456 ** 0.000	0.622 ** 0.000
Video is blocky	Corr. coeff. Sig (2-tailed)	0.272 * 0.006	0.540 ** 0.000	0.495 ** 0.000	0.684 ** 0.000		0.643 ** 0.000	0.454 ** 0.000	0.534 ** 0.000
Video is interrupted	Corr. coeff. Sig (2-tailed)	0.197 * 0.051	0.349 * 0.000	0.477 ** 0.000	0.451 ** 0.000	0.423** 0.000		0.649 ** 0.000	0.592 ** 0.000
I get disconnected	Corr. coeff. Sig (2-tailed)	0.117 * 0.247	0.288 * 0.004	0.443 ** 0.000	0.324 * 0.001	0.429** 0.000	0.463 ** 0.000		0.718 *** 0.000
Others get disconnected	Corr. coeff. Sig (2-tailed)	0.295 * 0.003	0.259 * 0.01	0.379 * 0.000	0.436 ** 0.000	0.426** 0.000	0.477 ** 0.000	0.706*** 0.000	

*** Strong, ** moderate, and * weak correlation based on the intervals found in [31].

From Table 8, we can conclude that all correlation coefficients $r_{frustration}$ are positive, which means that the two variables in a pair vary in the same direction. Most correlation coefficients are in the interval [0.4, 0.69], which means we found a moderate association of the ranks (**). Five correlation coefficients indicate a strong correlation (***) between the variables in a pair. Moderate and strong correlations show that participants who were frustrated by one type of degradation were also frustrated by another type. If they were not frustrated by one type of degradation, they were not frustrated by another type.

We can also observe how the correlation coefficient values are higher between pairs of variables that belong to the same degradation group (audio- or video-related degradations or meeting disconnections). Participants sensitive to a specific type of audio degradation are also susceptible to another one from that group. For instance, for echo in the audio-high-pitch audio pair $r_{frustration} = 0.809$ or high-pitch audio-audio contains noise pair where $r_{frustration} = 0.740$. The same applies to video-related degradations. For video is blurred-video is blocky pair $r_{frustration} = 0.848$. The trend is also observed for meeting disconnections ($r_{frustration} = 0.718$ for the I get disconnected-others get disconnected pair).

If we compare the correlation coefficients for the level of frustration and the frequency of degradation, it can be noticed that $r_{frustration}$ is generally higher for those pairs for which $r_{frequency}$ is also higher (see the values of the coefficients in the blue and red rectangles). This reaffirms the hypothesis that the frustration level is related to the degradation frequency but does not only depend it. We can see a lower association level between pairs of frequency of degradation than pairs of frustration with degradation (moderate and weak correlation).

A weaker association can be observed between frustration with degradation (moderate correlation) for pairs that do not belong to the same degradation group. It is important to note that the correlation is predominantly weak or moderate in some cases between the same pairs of degradation frequencies, which again confirms the hypothesis that the degradation frequency does not necessarily affect the level of frustration. Some coefficients are not even statistically significant ($p > 0.05$, $r_{frequency} = 0.197$ for the video is interrupted -echo in the audio pair and $r_{frequency} = 0.117$ for the I get disconnected-echo in the audio pair).

Table 8 also shows the p-values for the correlation test. The values presented are mainly <0.0001, showing that the correlations are significant at the 0.05 and the 0.01 level (two-tailed). The two-tailed confidence interval of correlation coefficient at 95% and 99% confidence levels are [−0.19755, 0.19755] and [−0.25776, 0.25776], respectively. Only the two previously mentioned coefficients do not meet these conditions.

4.4. User Frustration Susceptibility Index (UFSI)

To improve the models' accuracy (shown in Section 4.2), while taking into account that participants sensitive to a specific type of degradation are also susceptible to another from the same group (show in the previous chapter), we developed a User Frustration Susceptibility Index (UFSI). The UFSI is based on the residual intervals. We follow the earlier example from Figure 3 and in Figure 4 depicting the residual plots for the same three degradation types. As seen from Figure 4, we have unwanted patterns, i.e., we cannot accept the regression coefficients (Equations (4)–(6)) for modeling purposes.

The calculated coefficients (Equations (4)–(6)) differ in the value representing the mean change in the response given a one-unit change in the predictor and by the coefficient sign, which indicates the direction of the relationship between a predictor variable and the response variable. Although some coefficients contribute very little to the model (they are not statistically significant), and, in turn, the corresponding variables can be excluded, we decided to continue working with all variables. The most similar value of the coefficient is for the predictor variable degradation frequency.

Figure 4 shows that the residual values are in the [−2.26, 2.34] interval for "echo in the audio", and [−2.28, 2.45] and [−2.43, 2.02] for video is blocky, and I get disconnected degradation, respectively. This represents a significant error since the participants' level of frustration is in [1,5] interval. Moreover, the fitted values do not exceed 4.03 for echo in the audio, even though some participants stated that they were frustrated to the max by it. The modeled level of frustration for those participants is within [2.6, 3.5] interval. On the other side of the plot, the minimal fitted value for echo in the audio is 1.84 (which interprets as I can be a little frustrated); still, some participants expressed No frustration whatsoever, i.e., their LoF equaled 1. The same discrepancies can be identified for other degradation types.

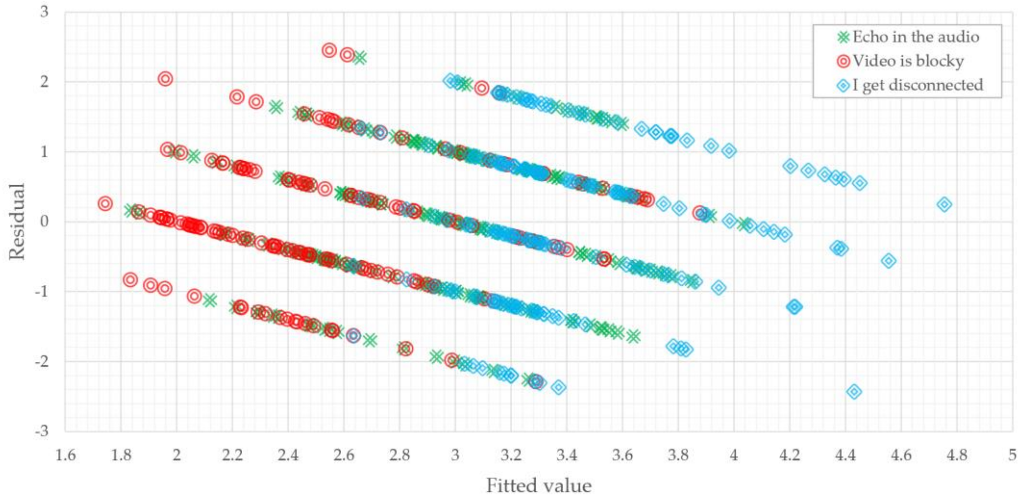

Figure 4. Residual plots for the three quality degradation types.

$$LoF_{a_{echo}} = 0.483 - 0.189 \cdot X_p + 0.034 \cdot X_r + 0.159 \cdot\ X_a + 0.029 \cdot X_m + 0.305 \cdot X_e - 0.007 \cdot X_{app} - 0.027 \cdot X_n + 0.524 \cdot X_f \quad (4)$$

$$LoF_{v_{blocky}} = -0.664 - 0.128 \cdot X_p + 0.078 \cdot X_r + 0.193 \cdot\ X_v + 0.120 \cdot\ X_s + 0.046 \cdot X_m - 0.020 \cdot X_e + 0.041 \cdot X_{app} + 0.102 \cdot X_n + 0.503 \cdot X_f \quad (5)$$

$$LoF_{d_{I\ get\ dis.}} = 2.342 + 0.028 \cdot X_p - 0.030 \cdot X_r - 0.051 \cdot X_m + 0.134 \cdot X_e - 0.011 \cdot X_{app} - 0.106 \cdot X_n + 0.567 \cdot X_f \quad (6)$$

As announced earlier, these residual analyses served as an incentive for developing UFSI. We developed UFSI to categorize users into groups based on their tendency to become frustrated with the impaired service. Figure 5 depicts the UFSI values that are based on the residual intervals and have linguistic meanings.

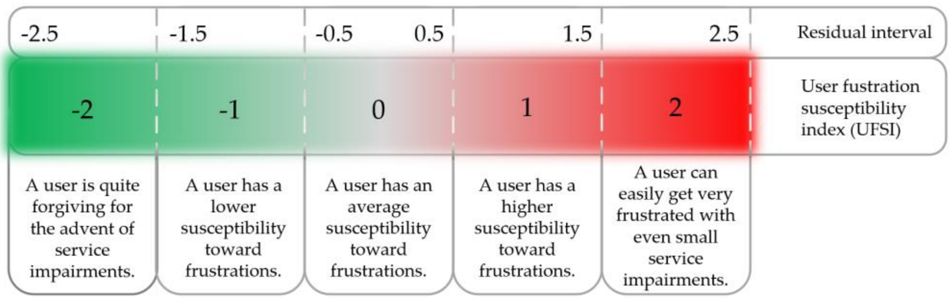

Figure 5. User frustration susceptibility index.

The participants were categorized as follows: The modeled level of frustration was subtracted from the participants' frustration level. The obtained values were divided into five intervals of equal length. Hence, participants whose level of frustration, for instance, was significantly higher than the expected level of frustration (given the values of the predictor variables) are classified in the A user can easily get very frustrated with even small service impairments group.

The UFSI becomes a new predictor variable in the models (Equations (1)–(3)). Using the same methodology described earlier (Section 4.2), we calculated the correlation coefficient and regression coefficients using matrix techniques and Excel's multiple regression analysis.

4.5. Multiple Correlation Coefficients and Updated Regression Model

After including UFSI into the models, new linear regression coefficients, standard errors, multiple correlation coefficients, and additional regression statistics are calculated and presented in Tables 9–11. As can be seen from the data shown in the tables, the correlation between these predictor variables and the LoF can be estimated as very strong. All multiple correlation coefficients are >0.9; R squared values are increased (from 12 and 22% to 92 and 96%), indicating the predictor variables now explain a significantly higher percentage of participants' frustration level variability. New values were also obtained for the ANOVA indicators; all *p*-values are less than 1.00E-05, which is less than 0.05 = α.

Table 9. The summary output for the audio group of degradations.

Regression Statistics	Echo in the Audio	High-Pitch Audio	Audio Contains Noise
Multiple R	0.96607	0.96739	0.95832
R Square	0.93329	0.93584	0.91839
Standard Error	0.28863	0.27422	0.30472
Observations	206	206	206

Table 10. The summary output for the video group of degradations.

Regression Statistics	Video Is Blurred	Video Is Blocky	Video Is Interrupted
Multiple R	0.96819	0.96344	0.96319
R Square	0.93740	0.92823	0.92773
Standard Error	0.25589	0.28659	0.30946
Observations	145	145	145

Table 11. The summary output for the disconnection group of degradations.

Regression Statistics	I Get Disconnected	Others Get Disconnected
Multiple R	0.98169	0.97097
R Square	0.96371	0.94279
Standard Error	0.23209	0.28933
Observations	175	215

Updated models with new values of regression coefficients estimate the LoF value more accurately, which is to be expected considering that we used the model assessment error to develop UFSI. For comparison reasons, Figure 6 shows the participants' frustration levels, degradation frequency, and new modeled frustration levels for three quality degradation types, as seen earlier in Figure 3. Figure 6 depicts that by introducing the UFSI, the model accurately estimates the user LoF for specific degradation. The modeled frustration levels are now grouped around the blue line (actual frustration levels) rather than, as seen earlier in Figure 3, around the gray dashed line showing degradation frequency. We reaffirm that with the residual analysis (Figure 7 shows the analysis results for the three degradations types as earlier). The residual values for all degradations that we analyzed are now <0.77. For a minor share of participants, the modeled frustration levels are inaccurate by more than 0.5. For echo in the audio, the video is blocky, and I get disconnected degradation that happens in 12, 8, and 6 test cases, respectively. Unlike before, the updated models can rate the frustration levels within the whole interval (i.e., from 1 to 5).

Tables 12–14 show the results obtained after testing the significance of regression coefficients to detect whether they differ from 0 or contribute to the model (for the three degradations for which we have presented the results in this paper). As in Section 4.4, we can observe that the coefficients differ in value and direction per each degradation type. The coefficients are approximately equal for predictor variables UFSI and degradation frequency (column coefficients in the tables). Note that the *p*-values for most of the coefficients are lower than 0.05 (the values written in bold text format in the *p*-value column). That means we cannot reject the hypothesis that they are 0 (i.e., they can be eliminated from the model). This is also confirmed since 0 lies in the interval between the lower 95% and upper 95% (i.e., the 95% confidence interval) for each of these coefficients (lower 95% and upper 95% columns in the tables). Although we would expect that the meeting role affects the user frustration levels when they are disconnected from the meeting, our data show otherwise.

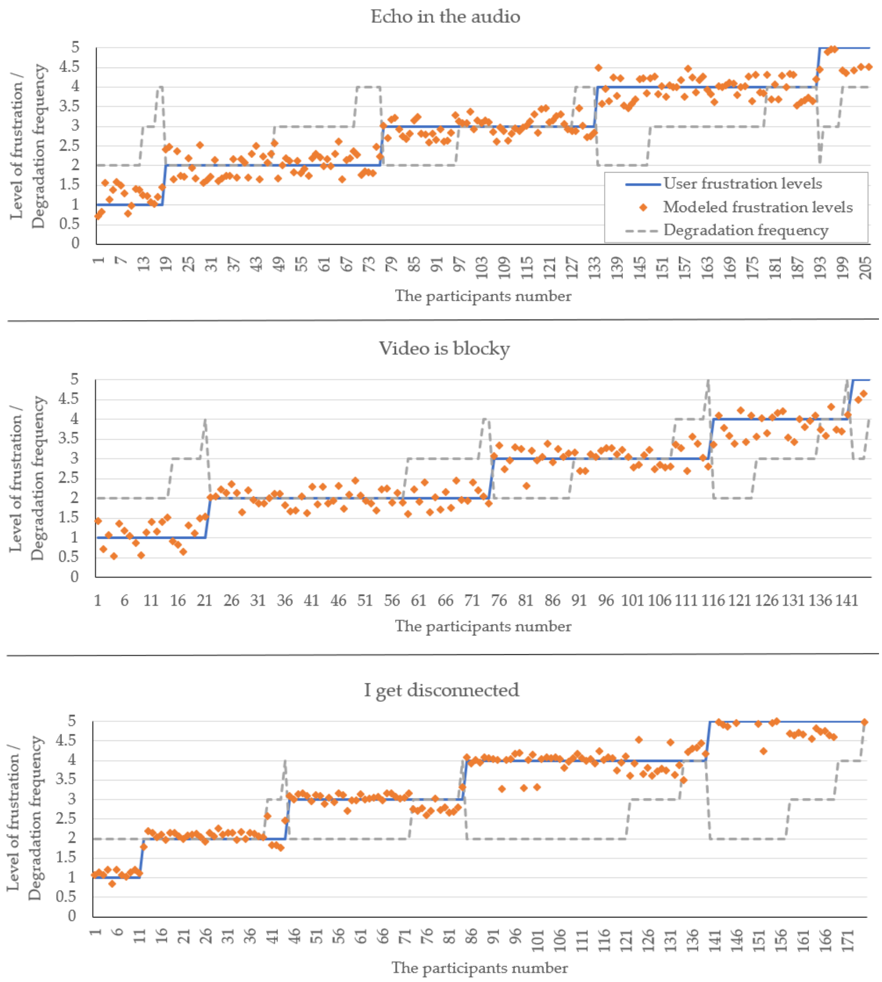

Figure 6. Participants' frustration levels, degradation frequency, and updated modeled frustration levels (after introducing UFSI) for three quality degradation types.

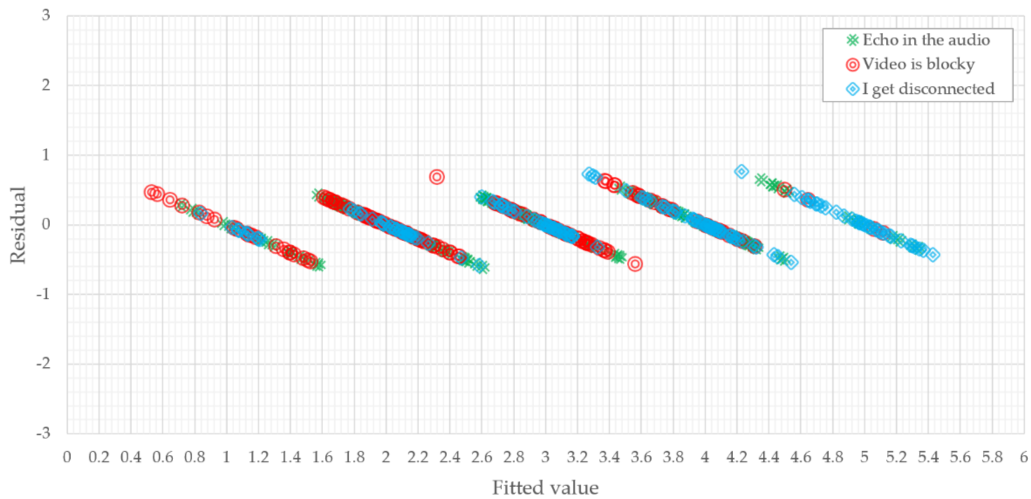

Figure 7. Updated residual plots for the three quality degradation types.

Table 12. Regression coefficients analysis for echo in the audio degradation.

Predictor Variable	Coefficients	Standard Error	t Stat.	*p*-Value	Lower 95%	Upper 95%
Intercept (β_0)	0.40648	0.28239	1.43941	0.15163	−0.15044	0.96340
Telemeeting purpose	−0.24394	0.03336	−7.31141	**0.00000**	−0.30974	−0.17814
Role on a telemeeting	0.06378	0.01337	4.76890	**0.00000**	0.03740	0.09016
Importance of the audio quality	0.18450	0.04976	3.70735	**0.00027**	0.08635	0.28264
Number of attended meetings (in one month)	0.03597	0.01957	1.83811	0.06756	−0.00262	0.07456
Equipment used	0.31806	0.03272	9.71996	**0.00000**	0.25353	0.38259
Application used	−0.00410	0.00529	−0.77476	0.43942	−0.01453	0.00633
Type of network	−0.06739	0.01897	−3.55260	**0.00048**	−0.10480	−0.02998
UFSI	0.90700	0.01912	47.43042	**0.00000**	0.86929	0.94471
Degradation frequency	0.53618	0.02851	18.80395	**0.00000**	0.47995	0.59241

Table 13. Regression coefficients analysis for video is blocky degradation.

Predictor Variable	Coefficients	Standard Error	t Stat.	*p*-Value	Lower 95%	Upper 95%
Intercept (β_0)	−0.81760	0.25052	−3.26357	**0.00139**	−1.31309	−0.32211
Telemeeting purpose	−0.13245	0.03724	−3.55635	**0.00052**	−0.20611	−0.05879
Role on a telemeeting	0.08629	0.01523	5.66707	**0.00000**	0.05617	0.11640
Share screen quality importance	0.07683	0.03781	2.03198	**0.04412**	0.00205	0.15161
Webcam video stream quality importance	0.18640	0.03053	6.10627	**0.00000**	0.12603	0.24678
Number of attended meetings (in one month)	0.07787	0.02269	3.43244	**0.00079**	0.03300	0.12274
Equipment used	0.02109	0.03879	0.54379	0.58748	−0.05562	0.09781
Application used	0.04001	0.00632	6.32707	**0.00000**	0.02750	0.05251
Type of network	0.04297	0.02352	1.82717	0.06988	−0.00354	0.08948
UFSI	0.94918	0.02614	36.31572	**0.00000**	0.89749	1.00088
Degradation frequency	0.60175	0.00843	17.85544	**0.00000**	0.53510	0.66841

Table 14. Regression coefficients analysis for I get disconnected degradation.

Predictor Variable	Coefficients	Standard Error	t Stat.	*p*-Value	Lower 95%	Upper 95%
Intercept (β_0)	1.84893	0.14025	13.18302	**0.00000**	1.57202	2.12583
Telemeeting purpose	0.03866	0.02715	1.42420	0.15627	−0.01494	0.09226
Role on a telemeeting	−0.01491	0.01113	−1.33973	0.18216	−0.03689	0.00707
Number of attended meetings (in one month)	−0.02107	0.01661	−1.26816	0.20652	−0.05387	0.01173
Equipment used	0.07977	0.02792	2.85753	**0.00482**	0.02466	0.13489
Application used	−0.01068	0.00449	−2.37821	**0.01853**	−0.01954	−0.00181
Type of network	−0.05418	0.01519	−3.56704	**0.00047**	−0.08416	−0.02419
UFSI	0.96693	0.01553	62.26445	**0.00000**	0.93627	0.99759
Degradation frequency	0.64708	0.02721	23.77757	**0.00000**	0.59335	0.70081

The same procedure was conducted for other degradation types, and that led us to the following conclusions.

- The values of the regression coefficients differ both in value and direction for different types of degradation and different degradation groups, showing how users react differently to various degradation types. This confirms the assumption that the level of frustration is influenced by the tendency of a person to become frustrated, taking into account that the values of regression coefficients shown in Tables 12–14 were obtained for different participants groups (Table 3);
- The values of the regression coefficients differ both in value and direction for different types of degradation within the same group, showing that users react differently to various degradations;
- Different scenarios were obtained for predictor variables that could be excluded from the model for different degradations since they do not significantly improve accuracy. Thus, we can observe that for: high-pitch audio degradation, all predictor variables defined in Section 4.2 are

significant;

- ○ audio contains noise degradation, predictor variables how important the audio quality is (X_a) and the number of meetings (X_m) can be omitted from the model;
- ○ video is blurred degradation, share screen quality (X_s) can be excluded as a predictor variable;
- ○ video is interrupted degradation multiple predictor variables can be excluded from the model, namely telemeeting purpose (X_p), equipment used in a telemeeting (X_e), and the importance of webcam video stream quality (X_v);
- ○ others get disconnected degradation, telemeeting purpose (X_p) as a predictor variable, can be omitted from the model.

4.6. Model Validation

As discussed in Section 4.1, the data collected with the survey were divided into the training set (T_r) and test set (T_e). So far, we have used the data from T_r for the model development. In this chapter, we employ T_e to validate the results from the models and justify the introduction of UFSI as a predictor variable.

To this end, we followed the same procedure as presented earlier. First, we introduced the predictor variable values in the corresponding models (Equations (1)–(3)) and determined the residuals. Second, based on the residuals, we mapped the UFSI as described in Figure 5. These values were then inserted into the models that include the UFSI as a predictor variable. The modeled LoF for the three degradations commented so far (namely, echo in the audio, video is blocky, and I get disconnected) are depicted in Figure 8. To determine how accurately the model predicts the LoF values, we use a frequently used measure of the root-mean-square error (RMSE) [32]. The RMSE values can be found in Table 15.

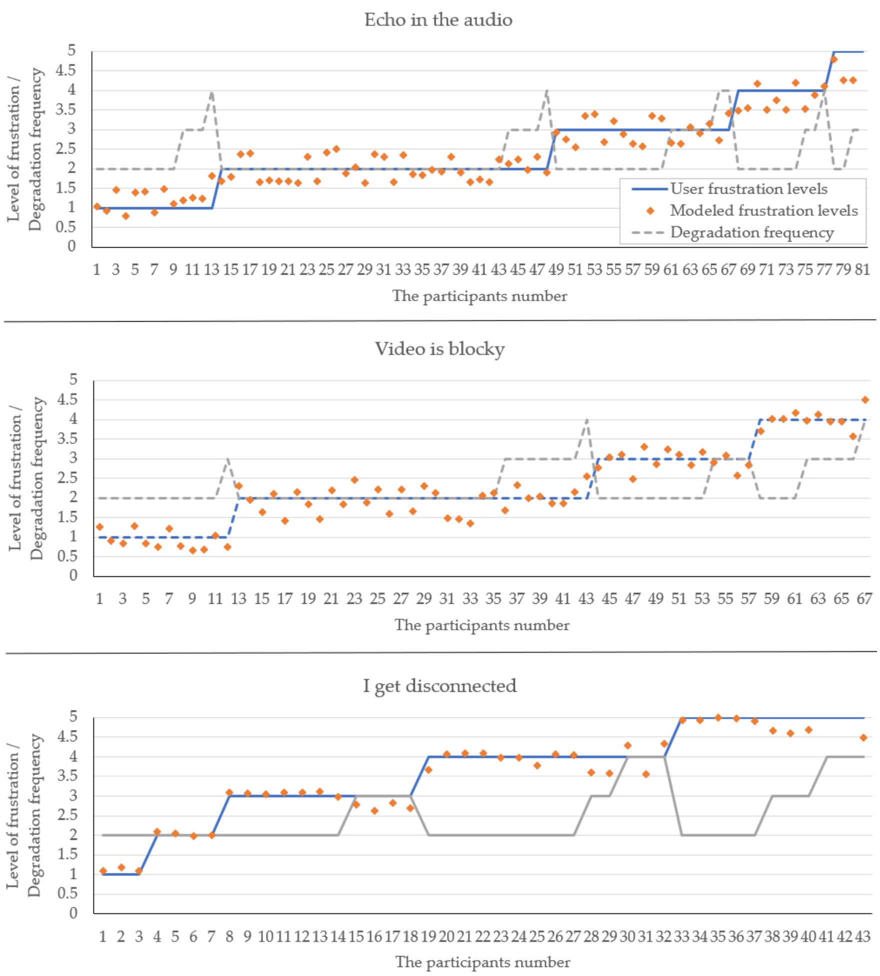

Figure 8. Using the test data sets, participants' frustration levels, degradation frequency, and modeled frustration levels (with UFSI) for three quality degradation types.

Table 15. RMSE values for the two data sets.

	RMSE Training Set	RMSE Test Set
Echo in the audio	0.282	0.334
High-pitch audio	0.227	0.348
Audio contains noise	0.192	0.329
Video is blurred	0.165	0.286
Video is blocky	0.276	0.283
Video is interrupted	0.203	0.243
I get disconnected	0.226	0.223
Others get disconnected	0.140	0.284

There is no fixed threshold for RMSE; it needs to be as small as possible and depends on the data set range a researcher is working with. From the data shown in Table 15, it can be concluded that models developed with UFSI as a predictor variable assess the user LoF accurately. We can see that all RMSE values between the two sets are similar; they are lower for the training set, except for the degradation I get disconnected. The largest difference between these values is obtained for the others becoming disconnected degradation type.

5. Conclusions

With this work, we have extended our understanding of how different quality degradations during telemeetings are related and how they affect users. The paper complements other research (e.g., the discoveries presented by Ammar et al. in [14]), which focused on predicting different audiovisual disturbances from network data flows. Our findings show how those disturbances may affect users. We showed how users aggravated by one type of degradation may become more susceptible to others, especially if they belong to the same group (audio- or video-related degradations or telemeeting disconnections).

The paper focused on modeling user frustration levels with the telemeeting service, which is hampered by the quality impairments (the impact of eight degradation types was analyzed). While examining the interplay between each degradation type and the associated level of participants' frustration, we found that some qualitative and quantitative factors that we investigated affect users, yet the effect was not as strong as anticipated (e.g., the impact of degradation frequency on user perception). This prompted us to look for other factors that might interfere with the user perception. At the end of this path was UFSI, which we developed to categorize users into groups based on their tendency to get frustrated with the impaired service. After we included UFSI as the predictor variable in the regression models, the modeled levels of user frustration became accurate in both training and test data set (validated with the RMSE values).

We will continue researching the factors affecting the user experience for this type of service and work on the model improvement in future work. This could lead us, for instance, to new data sets that will be formed based on session-specific user responses and QoS.

Author Contributions: Conceptualization, M.M. and Š.M.; methodology, Š.M. and M.M.; validation, Š.M.; formal analysis, Š.M.; investigation, M.M. and Š.M.; resources, M.M.; data curation, M.M.; writing—original draft preparation, M.M.; writing—review and editing, Š.M.; visualization, Š.M.; supervision, Š.M. All authors have read and agreed to the published version of the manuscript.

Funding: The publication of this research was funded by the University of Zagreb under Grants for financing scientific and artistic activities of the University of Zagreb in the 2020/2021 program (internal project id: 210219 ZUID 2020/2021).

Institutional Review Board Statement: This study's ethical review and approval were waived since the authors did not conduct any experiments with the participants. Participation was included only the completion of the questionnaire, and it was voluntary.

Informed Consent Statement: The survey participants' consent was collected before starting the online questionnaire. The participants were informed that their identity would remain unknown to the authors; no personal data were collected with the survey.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors wish to thank Borna Abramović for promoting the survey on online platforms and social media and assisting in the initial data analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. GrandViewResearch. Available online: https://www.grandviewresearch.com/industry-analysis/video-conferencing-market (accessed on 5 April 2021).
2. Gminsights. Available online: https://www.gminsights.com/industry-analysis/video-conferencing-market (accessed on 5 April 2021).
3. Matulin, M.; Mrvelj, Š. Modelling user quality of experience from objective and subjective data sets using fuzzy logic. *Multimed. Syst.* **2018**, *24*, 645–667. [CrossRef]
4. Qualinet. Available online: http://www.qualinet.eu/images/stories/QoE_whitepaper_v1.2.pdf (accessed on 4 April 2021).
5. ITU-T. Subjective Quality Evaluation of Audio and Audiovisual Multiparty Telemeetings. *Document ITU-T Rec. P.1301.* 2017. Available online: https://www.itu.int/rec/T-REC-P.1301-201710-I/en (accessed on 17 April 2021).
6. Bakar, G.; Kirmizioglu, R.A.; Tekalp, A.M. Motion-based rate adaptation in WebRTC videoconferencing using scalable video coding. *IEEE Trans. Multimed.* **2019**, *21*, 429–441. [CrossRef]
7. 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. Multimed.* **2013**, *15*, 1638–1652. [CrossRef]
8. Li, C.L.; Xiong, H.K.; Zou, J.N.; Wu, D.P.O. Joint dynamic rate control and transmission scheduling for scalable video multirate multicast over wireless networks. *IEEE Trans. Multimed.* **2018**, *20*, 361–378. [CrossRef]
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.* **2018**, *132*, 42–55. [CrossRef]
10. Hoßfeld, T.; Heegaard, P.E.; Varela, M.; Skorin-Kapov, L.; Fiedler, M. From QoS Distributions to QoE Distributions: A System's Perspective. In Proceedings of the 6th IEEE Conference on Network Softwarization (NetSoft), Ghent, Belgium, 29 June–3 July 2020; IEEE: New York, NY, USA; pp. 51–56. [CrossRef]
11. Belmudez, B. *Audiovisual Quality Assessment and Prediction for Videotelephony*, 1st ed.; Springer International Publishing: Berlin, Germany, 2015.
12. Laghari, K.U.R.; 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, CA, USA, 3–6 December 2012; IEEE: New York, NY, USA; pp. 1–9.
13. 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 Proceedings of the 15th International Wireless Communications & Mobile Computing Conference (IWCMC), Tangier, Morocco, 24–28 June 2019; IEEE: New York, NY, USA; pp. 1267–1272. [CrossRef]
14. Ammar, D.; Moor, K.D.; Skorin-Kapov, L.; Fiedler, M.; Heegaard, P.E. Exploring the usefulness of machine learning in the context of WebRTC performance estimation. In Proceedings of the 44th Conference on Local Computer Networks (LCN), Osnabrueck, Germany, 13 February 2019; IEEE: New York, NY, USA; pp. 406–413. [CrossRef]
15. Vučić, D.; Skorin-Kapov, L. The impact of mobile device factors on QoE for multi-party video conferencing via WebRTC. In Proceedings of the 13th International Conference on Telecommunications (ConTEL), Graz, Austria, 13–15 July 2015; IEEE: New York, NY, USA; pp. 1–8. [CrossRef]
16. Vučić, D.; Skorin-Kapov, L.; Sužnjević, M. The impact of bandwidth limitations and video resolution size on QoE for WebRTC-based mobile multi-party video conferencing. In Proceedings of the 5th ISCA/DEGA Workshop on Perceptual Quality of Systems (PQS 2016), Berlin, Germany, 29–31 August 2016; ISCA: Grenoble, France; pp. 59–63. [CrossRef]
17. Vučić, D.; Skorin-Kapov, L. The impact of packet loss and Google congestion control on QoE for WebRTC-based mobile multiparty audiovisual telemeetings. *MultiMedia Modeling Lect. Notes Comput. Sci.* **2019**, *11295*, 459–470. [CrossRef]
18. Vučić, D.; Skorin-Kapov, L. QoE evaluation of WebRTC-based mobile multiparty video calls in light of different video codec settings. In Proceedings of the 15th International Conference on Telecommunications (ConTEL), Graz, Austria, 3–5 July 2019; IEEE: New York, NY, USA; pp. 1–8. [CrossRef]
19. Vučić, D.; Skorin-Kapov, L. QoE Assessment of Mobile Multiparty Audiovisual Telemeetings. *IEEE Access* **2020**, *8*, 107669–107684. [CrossRef]
20. Schmitt, M.; Redi, J.; Bulterman, D.; Cesar, P.S. Towards individual QoE for multiparty videoconferencing. *IEEE Trans. Multimed.* **2018**, *20*, 1781–1795. [CrossRef]
21. Schmitt, M.; Redi, J.; Cesar, P.; Bulterman, D. 1Mbps is enough: Video quality and individual idiosyncrasies in multiparty HD videoconferencing. In Proceedings of the 8th International Conference on Quality of Multimedia Experience (QoMEX), Lisbon, Portugal, 6–8 June 2016; IEEE: New York, NY, USA; pp. 1–6. [CrossRef]
22. De Moor, K.; Arndt, S.; Ammar, D.; Voigt-Antons, J.-N.; Perkis, A.; Heegaard, P.E. Exploring diverse measures for evaluating QoE in the context of WebRTC. In Proceedings of the 9th International Conference on Quality of Multimedia Experience (QoMEX), Erfurt, Germany, 31 May–2 June 2017; IEEE: New York, NY, USA; pp. 1–3. [CrossRef]

23. Husić, J.B.; Baraković, S.; Veispahić, A. What factors influence the quality of experience for WebRTC video calls? In Proceedings of the 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 22–26 May 2017; IEEE: New York, NY, USA; pp. 428–433. [CrossRef]
24. Silva, A.F.D.; Mylène, C.Q. Perceptual strengths of video impairments that combine blockiness, blurriness, and packet-loss artifacts. In Proceedings of the International Symposium on Electronic Imaging, San Francisco, CA, USA, 28 January–1 February 2018; IS&T: Washington, DC, USA; pp. 2341–2346. [CrossRef]
25. Øie, E.B.; Koniuch, K.; Cieplińska, N.; De Moor, K. Factors influencing QoE of video consultations. In Proceedings of the 13th International Conference on Quality of Multimedia Experience (QoMEX), Montreal, QC, Canada, 14–17 June 2021; IEEE: New York, NY, USA; pp. 137–140. [CrossRef]
26. Matulin, M.; Mrvelj, Š.; Abramović, B. How frustrated are you? User perception about different videoconference quality degradations. In Proceedings of the 5th EAI International Conference on Management of Manufacturing Systems, Online, 27–28 October 2020. [CrossRef]
27. Matulin, M.; Mrvelj, Š.; Abramović, B.; Šoštarić, T.; Čejvan, M. User Quality of Experience comparison between Skype, Microsoft Teams and Zoom videoconferencing tools. In Proceedings of the 5th EAI International Conference on Future Access Enablers of Ubiquitous and Intelligent Infrastructures, Zagreb, Croatia (Online), 6–7 May 2021. [CrossRef]
28. Gao, Y.; Wei, X.; Zhou, L. Personalized QoE Improvement for Networking Video Service. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 2311–2323. [CrossRef]
29. Pezzulli, S.; Martini, M.G.; Barman, N. Estimation of Quality Scores from Subjective Tests-Beyond Subjects' MOS. *IEEE Trans. Multimed.* **2021**, *23*, 2505–2519. [CrossRef]
30. Mrvelj, Š.; Matulin, M. Impact of packet loss on the perceived quality of UDP-based multimedia streaming: A study of user quality of experience in real-life environments. *Multimed. Syst.* **2018**, *24*, 33–53. [CrossRef]
31. Schober, P.; Boer, C.; Schwarte, L.A. Correlation coefficients: Appropriate use and interpretation. *Anesth. Analg.* **2018**, *126*, 1763–1768. [CrossRef] [PubMed]
32. Hyndman, R.J.; Koehler, A.B. Another look at measures of forecast accuracy. *Int. J. Forecast.* **2006**, *22*, 679–688. [CrossRef]