Assessment of the Simulator Sickness Questionnaire for Omnidirectional Videos

Ashutosh Singla* Steve Göring† Dominik Keller‡ Rakesh Rao Ramachandra Rao§
Stephan Fremerey¶ Alexander Raake‖

Audiovisual Technology Group (AVT)
Technical University of Ilmenau, Germany

ABSTRACT

Virtual Reality/360° videos provide an immersive experience to users. Besides this, 360° videos may lead to an undesirable effect when consumed with Head-Mounted Displays (HMDs), referred to as simulator sickness/cybersickness. The Simulator Sickness Questionnaire (SSQ) is the most widely used questionnaire for the assessment of simulator sickness. Since the SSQ with its 16 questions was not designed for 360° video related studies, our research hypothesis in this paper was that it may be simplified to enable more efficient testing for 360° video. Hence, we evaluate the SSQ to reduce the number of questions asked from subjects, based on six different previously conducted studies. We derive the reduced set of questions from the SSQ using Principal Component Analysis (PCA) for each test. Pearson Correlation is analysed to compare the relation of all obtained reduced questionnaires as well as two further variants of SSQ reported in the literature, namely Virtual Reality Sickness Questionnaire (VRSQ) and Cybersickness Questionnaire (CSQ). Our analysis suggests that a reduced questionnaire with 9 out of 16 questions yields the best agreement with the initial SSQ, with less than 44% of the initial questions. Exploratory Factor Analysis (EFA) shows that the nine symptom-related attributes determined as relevant by PCA also appear to be sufficient to represent the three dimensions resulting from EFA, namely, Uneasiness, Visual Discomfort and Loss of Balance. The simplified version of the SSQ has the potential to be more efficiently used than the initial SSQ for 360° video by focusing on the questions that are most relevant for individuals, shortening the required testing time.

Index Terms: 360° video—simulator sickness—questionnaire—cybersickness; PCA—factor analysis

1 INTRODUCTION

Simulator Sickness (or cybersickness) is triggered by visual stimuli during exposure in a Virtual Environment (VE) via HMDs. Different symptoms can be induced during exposure such as headache, dizziness, eyestrain, blurred vision, etc. [6, 26]. Virtual Reality (VR) videos provide an immersive experience to users. Besides this, users may experience simulator sickness which may affect their comfort and viewing experience. To efficiently and accurately self-report the symptoms, an appropriate test method is required. The most popular questionnaire used is the SSQ developed by Kennedy et al. in 1993 [18]. This SSQ was designed mainly for measuring simulator sickness experienced by Navy pilots for flight simulator studies. There are differences between simulator and VR systems and it may be possible that some of the symptoms that are relevant for simulator-system studies may not be relevant for VR, and in particular for 360° video. Furthermore, some symptoms belong to more than one category. The SSQ was adapted for VR contents by [11, 20], but not for 360° contents. VR and 360° contents differ in terms of Degrees of Freedom (DoF) and the way they are captured. Here, VR video contains mainly computer-generated imagery and can provide movement to a user in 6 DoF (3 directional and 3 rotational). On the other hand, 360° video is captured from a real-life scene using 360° camera and can provide movement to a user in only 3 DoF (3 rotational). Hence, it could be possible that users may feel less prone to simulator sickness in 360° videos than VR videos. Therefore, it would be reasonable to investigate if it is efficient to use the SSQ in its current form for 360° videos [11, 23]. Along similar lines, there arises the question of whether the set of 16 questions may be further reduced when going from a simulator to 360° video, to enable a more dedicated and time-efficient assessment. The resulting simplified version of the SSQ focuses on the symptoms that are relevant for 360° video related studies, with the aim to reduce test time and ultimately enable larger-scale testing. Please note that we are not arguing that the SSQ cannot be used in the case of 360° videos. It may be possible that individual symptoms are less relevant for 360° video.

We analyzed the data of our previously published papers [27, 28, 30, 31] in more depth to derive different sets of simplified questionnaires from the SSQ. The main contributions of this paper are as follows: (1) proposing a simplified SSQ for the self-reporting of symptoms which are relevant for 360° videos, without excluding any relevant symptoms, (2) comparing different versions of simplified SSQs with the state-of-the-art questionnaires, CSQ and VRSQ and with the initial SSQ set of scales, and (3) a simulator sickness database (158 subjects) with subjectively annotated raw scores for different 360° videos and sessions, generated using a wide variety of sequences, different HMDs, different durations of test sessions and with different test designs.

The paper is organized as follows: Section 2 briefly describes related work. The experimental setup, 360° video dataset and test methods are described in Section 3 for all the subjective tests. Section 4 presents and discusses the experimental results. Concluding remarks are given in Section 5.

*e-mail: ashutosh.singla@tu-ilmenau.de
†e-mail:steve.goering@tu-ilmenau.de
‡e-mail:dominik.keller@tu-ilmenau.de
§e-mail:rakesh-rao.ramachandra-rao@tu-ilmenau.de
¶e-mail:stephan.fremerey@tu-ilmenau.de
‖e-mail:alexander.raake@tu-ilmenau.de

2642-5254/21/$31.00 ©2021 IEEE
DOI 10.1109/VR50410.2021.00041

2 Related Work

The extent of simulator sickness symptoms experienced by a user can be measured by using objective and subjective methods [37]. Objective measures include measurement of physiological signals such as heart rate, galvanic skin response, and respiration [7, 15]. While subjective measures include self-reporting of symptoms using long questionnaires such as the SSQ [18], VRSQ [20], CSQ [11], MASQ (Motion Sickness Assessment Questionnaire) [9], or using rapid or fast self-report methods (Short Questionnaire or single scale question) e.g., Fast Motion Sickness Scale (FMS) [19], Misery Scale (MISC) [38], Short SSQ [36] and comfort scale [22]. Furthermore, simulator sickness can be assessed by measuring postural instability [21, 33].

Kennedy et al. developed the SSQ [18] based on the Motion Sickness Questionnaire (MSQ) [17]. They removed those symptoms that occur infrequently and may give misleading indications. Twelve symptoms were eliminated out of 28, resulting in 16 symptoms. These 16 symptoms are further divided into three factors Nausea, Oculomotor and Disorientation. Kim et al. attempted to reduce the number of questions by using PCA and proposed VRSQ [20] that contains only 9 symptoms, which are divided into two factors Oculomotor and Disorientation. However, there are some limitations of VRSQ, since other research has shown that deleting the "Nausea" component from the SSQ may not be appropriate. For example, the studies by Singla et al. [28] and Rebenitsch et al. [23] indicate that "Nausea" was experienced in their tests for VR/360° videos. Therefore, "Nausea", a sub-category of SSQ, should not be completely removed from the corresponding test questions. Ill [11] derived 9 symptoms from the SSQ and proposed a two-factor structure: Dizziness and Difficulty in Focusing. However, they used a very limited dataset consisting of only three VR games and one HMD (HTC Vive) to derive the 9 symptoms. Bouchard et al. [4] proposed to re-factor the three-factor structure for the SSQ to the two-factor structure (Nausea and Oculomotor) without removing any of the symptoms from the SSQ. They pointed out that the three-factor structure of the SSQ was administered on military participants using simulators, and that these factors may not be suitable for research related to virtual reality. The main limitation of their work in light of the goals of the present work is that they have not reduced the number of symptoms.

The single-scale questions are often used to measure the amount of simulator sickness using a Likert like scale [19, 36, 38]. The FMS was proposed by Keshavarz et al. [19]. The scale used ranges from 0 to 20, where 0 is no sickness at all and 20 means frank sickness i.e. imminent vomiting. The MISC was developed by Wertheim et al. [38] to assess the overall well-being of the person, using a 0-10 scale (0: no problems, 1: uneasy feeling in head ... 10: vomiting). Tran et al. [36] proposed the short SSQ that assesses the level of nausea or dizziness on a 5-point scale (1: very dizzy, 2: dizzy, 3: slightly dizzy, 4: not dizzy, and 5: absolutely not dizzy). All simulator sickness assessment methods described above used discrete scales.

In [29], Singla et al. compared a Short SSQ with the widely used SSQ [18] (Long SSQ) for the case of viewing 360° videos. Their results indicate that the Short SSQ cannot replace the Long SSQ to evaluate the comparatively lower impact of individual technical factors such as bitrate, resolution, etc., but can effectively replace the long SSQ to distinguish between video contents, for example, to assess whether they produce high or low degrees of simulator sickness [29].

Despite having multiple long questionnaires and single-scale questions, all of them come along with certain limitations, especially when a larger-scale test-data collection is sought. On the one hand, as pointed out above, certain single-scale questions cannot fully record the symptoms. On the other hand, the problem with the long questionnaire (SSQ) is that users may get bored while filling in too many questions after each video/session, and corresponding tests are time-consuming so that only rather few conditions can be tested in one session.

Our work contributes towards proposing a simplified SSQ for the self-reporting of symptoms which are relevant for 360° videos, without excluding any relevant symptoms. The database used for the development of this simplified SSQ includes more than 150 test participants, representing data from six different studies. Furthermore, we are using a wide variety of sequences, different HMDs, different durations of test sessions with different test designs and test methods to derive the most relevant symptoms.

Table 1: Information about different Tests.

Test #	HMD Name	Number of Subjects	Number of outliers	Number of times SSQ Filled	Duration of VR session
1 [28]	HTC Vive & Oculus Rift	28	0	24	1 min
2 [27]	Oculus Rift	30	1	4	15 mins
3 [30]	Oculus Rift	27	1	4	10 mins
4 [31]	HTC Vive Pro	29	4	4	7 mins
5 [31]	HTC Vive Pro	30	5	4	11 mins
6 [31]	HTC Vive Pro	28	3	4	11 mins

Table 2: Principal Component Analysis on raw simulator sickness scores for Subject_ID 1 for Test #1.

Variance	34.859	19.99	10.613	10.004	8.366
Symptom	PC 1	PC 2	PC 3	PC 4	PC 5
General Discomfort	0.588	0.082	-0.124	-0.342	-0.015
Fatigue	0.1	-0.493	-0.313	0.633	-0.086
Headache	0.239	-0.414	0.169	-0.132	0.761
Eye Strain	0.236	0.614	-0.437	0.393	0.358
Difficulty Focusing	0	0	0	0	0
Salivation Increasing	0	0	0	0	0
Sweating	0	0	0	0	0
Nausea	0.455	0.071	-0.097	-0.029	-0.109
Difficulty Concentrating	0.256	-0.396	-0.315	0	-0.297
Fullness of Head	0.306	0.038	0.679	0.485	0.038
Blurred Vision	0.017	0.193	0.243	0.174	-0.268
Dizziness with Eye Open	0.366	0.014	0.19	-0.073	-0.331
Dizziness with Eye Closed	0.123	0.017	0.037	-0.141	-0.035
Vertigo	0.123	0.017	0.037	-0.141	-0.035
Stomach Awareness	0	0	0	0	0
Burping	0	0	0	0	0

3 Experiment

For test 1, six different omnidirectional videos (1: Mega Coaster, 2: Project 360, 3: 360 Cockpit View, 4: Sky Diving in 360, 5: Reimagine – Etihad A380, 6: Surrounded by Wild Elephants) including audio were downloaded from YouTube[1] in two resolutions, namely Full HD and 4K, and no further encoding was performed. The downloaded video sequences are in YUV 4:2:0 colour space, 8-bits per colour channel with a duration of 60 s. Each subject had to rate 24 PVSs (Processed Video Sequences) (6 sequences, 2 resolutions and 2 HMDs) during the test. First, each subject had to watch the video for 60 s with a specific HMD. Then, the subject had 60 s to rate the integral quality of the presentation using the 5-point ACR test method (ITU-T P.910) and to fill in the SSQ [18]. At last, a 60 s break was given to recover from the adverse and cumulative

[1]These videos are not raw videos and have the YouTube specific encoding.

effect of one video after another on simulator sickness scores shown to be present for longer viewing sessions, see [22].

For the tests 2 and 3, six different omnidirectional videos without audio were downloaded from [2, 25, 34] (1: DrivingInCountry, 2: PoleVault_le, 3: Gaslamp, 4: Harbor, 5: KiteFlite, 6: Trolley). The video sequences are in YUV 4:2:0 color space, 8-bits per pixel with a duration of 10 s. For encoding the source video sequences, FFmpeg with libx265 was used. The Video Buffering Verifier (VBV) method was used for encoding, following a one-pass encoding scheme. Each subject had to rate 60 PVSs (6 sequences, 2 resolutions[2] and 5 bitrates) during the test. The whole test was divided into four test sessions. In each session, 15 360° videos were rated using the M–ACR (test 2, [27]) and DSIS (test 3, [13]) test methods. After the completion of each test session, subjects were asked to fill in the SSQ. The duration of each test session was 15 minutes for test 2 and 10 minutes for test 3. After every test session, the subjects were given a break of equal duration to that of the test session, also to recover from possible simulator sickness to some extent.

For the tests 4, 5 and 6, five different omnidirectional videos without audio were downloaded from [1, 2] (1: Gaslamp, 2: Harbor, 3: KiteFlite, 4: Skateboard_in_lot, 5: Trolley). The video sequences are in YUV 4:2:0 color space, 8-bits per pixel with a duration of 10 s. For encoding the source videos, FFmpeg with libx265 was applied. Two-pass (preset slow) encoding with the specified target bit-rate was used. Each subject had to rate 60 PVS (5 sequences, 3 resolutions[3] and 4 bitrates) during the test. The whole test was divided into four test sessions. In each session, 15 360° videos were rated using the ACR (test 4, [13]), M–ACR (test 5, [27]) and DSIS (test 6, [13]) test method. After the completion of each test session, subjects were asked to fill in the SSQ. The duration of each test session for tests 4, 5 and 6 was 7, 11 and 11 minutes, respectively. After every test session, subjects were given a break of equal duration to that of the test session.

For all the tests, the Whirligig player was used for the playout of the 360° videos. Table 1 lists the information about HMD names, number of participants and outliers, as well as the duration of the VR sessions. The outlier detection was performed on the quality ratings to check the reliability of the test participants, based on ITU-R Rec. BT.500-13 [12] for test 1, and the Pearson correlation coefficient [35] for tests 2–6. To find the suitable test method for the assessment of simulator sickness, we consider data only from reliable subjects. Accordingly, one participant had to be excluded in test 2, one in test 3, four in test 4, five in test 5, three in test 6 and no participant had to be excluded in test 1. Thus, we consider the response from 158 valid subjects after removing the outliers (see Table 1). All collected data is available online[4].

For the tests 1 and 2, the subjects were asked to fill in the SSQ after watching the stimuli (test 1) or after the test session (test 2). For the tests 3, 4, 5 and 6, the SSQ was filled after the pre-screening, thus before the training session. In addition, for the latter four tests, responses were given after each respective test session, when subjects had viewed a certain number (test dependent) of 360° videos. Note that using the SSQ at the beginning of a test was not done to screen e.g., for unhealthy participants or to bias their response. It was used with the intention only to measure their current sickness level as most of the symptoms can occur to a participant without the use of a simulator or an HMD. For the tests 4, 5 and 6, the SSQ was filled after the break as well to investigate the impact of breaks on the scores of simulator sickness. These responses were not considered for the development of the simplified questionnaire.

[2] 4K and FHD

[3] 4K, 6K and 8K

[4] https://github.com/Telecommunication-Telemedia-Assessment/360-SimulatorSickness-Data

4 Results

4.1 Per-subject PCA

PCA is a dimensionality reduction technique that is used to find the essential variables without losing any important information or properties and this reduces the complexity of the data [16,20,24]. PCA has been applied to the SSQ to reduce the number of symptoms [20] and to find the factor structure [5]. In addition, factor analysis is done to group the symptoms into different subgroups based on their loading value. In this work, we followed a similar procedure applied in the original paper by Kennedy et al. [18] for factor analysis.

The goal for applying PCA on a per-subject level is to find the most relevant symptoms for each subject and later analyze the number of subjects for whom a particular symptom is relevant. Based on the percentage of occurrence, we decided whether to include a given symptom in the final questionnaire for each test or not.

In order to keep the Pearson Correlation Coefficients (PCC) of sub-categories/clusters (Nausea (N), Oculomotor (O) and Disorientation (D)) and TotalScore (TS) for the original scoring and the simplified scoring above 95%[5], we selected the different thresholds for variance explained by the number of Principal Components (PCs), magnitude for symptom and % of people accordingly.

We apply a PCA on the raw scores of simulator sickness for each subject to get the principal component coefficients and limit the amount of total variance to at least 80% explained by the number of PCs. There is no objective way to find how much variance is needed or how many PCs are sufficient. Therefore, we selected 80% which is an acceptably large variance. Table 2 shows the exemplary principal component coefficients for subject ID_1. It can be seen that five PCs are needed to have a variance of at least 80% for subject ID_1. From Table 2 we can see that some of the symptoms such as Difficulty Focusing, Salivation Increasing, Sweating, Stomach Awareness and Burping have almost zero contribution to the PCs. The results further show that different numbers of PCs are needed for each subject, as we have fixed the variance not the number of PCs. So, for some subjects, we may need a higher or lower number of PCs. Table 4 shows the number of PCs required to have a variance of at least 80% and the obtained magnitude for each of the symptoms for each subject. We calculate the magnitude of each symptom across the selected PCs using Eq. 5. Here, J is the number of PCs needed to yield a variance of at least 80%. The obtained magnitude for each of the symptoms will be used for the selection of the most important symptoms.

$$Symptom_i = \sqrt{\sum_{n=1}^{J} PC\_n_i^2} \tag{5}$$

The two subjects (ID_11, and ID_27) did not report any of the symptoms, which is why the magnitude of each symptom for these subjects is zero. In a subsequent thresholding, symptoms with magnitude greater than 0.4 are taken as important for a subject and counted as relevant. Afterwards, we calculate for how many subjects (in %) a symptom is relevant, which is shown by the last column of Table 4. The symptoms such as

[5] We selected different thresholds and 95% comes out to be the best

$$N = (GeneralDiscomfort + IncreasedSalivation + Sweating + Nausea + DifficultyConcentrating + StomachAwareness + Burping) \times 9.54$$
$$O = (GeneralDiscomfort + Fatigue + Headache + EyeStrain + DifficultyFocusing + DifficultyConcentrating + BlurredVision) \times 7.58$$
$$D = (DifficultyFocusing + Nausea + FullnessOfHead + DizzinessWithEyeOpen + BlurredVision + DizzinessWithEyeClosed + Vertigo) \times 13.92$$
$$TS = \left(\frac{N}{9.54} + \frac{O}{7.58} + \frac{D}{13.92}\right) \times 3.74 \quad (1)$$

$$N = (GeneralDiscomfort + DifficultyConcentrating) \times 9.54$$
$$O = (GeneralDiscomfort + EyeStrain + DifficultyFocusing + DifficultyConcentrating + BlurredVision) \times 7.58$$
$$D = (DifficultyFocusing + DizzinessWithEyeOpen + BlurredVision) \times 13.92$$
$$TS = \left(\frac{N}{9.54} + \frac{O}{7.58} + \frac{D}{13.92}\right) \times 3.74 \quad (2)$$

Table 3: Variance and Magnitude Thresholds for different tests

Test #	Variance (in %)	Magnitude_Threshold	% of People
1	80	0.4	≥50
2	80	0.2	≥45
3	80	0.3	≥50
4	80	0.25	>40
5	80	0.2	>40
6	80	0.3	>40

Burping, Sweating and Stomach Awareness have occurred only for a very few subjects. To keep the simplified TS aligned to the original TS, symptoms being relevant for at least 50% of the subjects are marked as relevant symptoms (highlighted in Table 4). Using this criterion, the number of symptoms was reduced from 16 to 6 for test #1, and these 6 symptoms are General Discomfort, Eye Strain, Difficulty Focusing, Difficulty Concentrating, Blurred Vision and Dizziness with Eye open. These 6 symptoms are classified into three sub-categories (N, O and D), using the same scoring method as described in the original SSQ [18], see Eq. 1. We calculate the average N, O, D and TS over all subjects for all video sequences based on the Eq. 2, for test #1. We performed a correlation analysis based on the simplified scoring and the original scoring. Results show high PCC of 0.969, 0.991, 0.956 and 0.982 for all the three sub-categories of SSQ (N, O and D) and the Total Score respectively for test #1.

Similarly, we perform per-subject PCA on the other five subjective tests and perform all the above-mentioned steps. As described in Section 3, each test has a different test design and objective, therefore, different thresholds were selected for each test. Table 3 lists the different thresholds for magnitude and (%) of people for tests #1 to #6. We obtain a different simplified questionnaire referred to as R1 to R6 for test #1 to test #6, respectively, as shown in Table 5, and classified the symptoms according to the initial sub-categories of the SSQ (N, O and D). It can clearly be seen from Table 5 that different tests have resulted in different sets of relevant symptoms, except for test #6. The relevant symptoms for tests #5 and #6 are exactly the same. Further, we derive three additional questionnaires (R7, R8 and R9) using questionnaires R1 to R6. We derive R7 using the union operation on all the questionnaires R1 to R6 ($R7 = R1 \cup R2 \cup R3 \cup R4 \cup R5 \cup R6$), resulting in 9 symptoms. For R8, we perform the intersection operation and include only those symptoms that belong to at least five different questionnaires ($R8 = R1 \cap R2 \cap R3 \cap R4 \cap R5 \cap R6, where, Symptom_i \in at\ least\ 5\ of\ \{R1, R2, ..R6\}$), resulting in 5 symptoms. R9 is derived by performing a union operation between R8 and Dizziness With Eye Open ($R9 = R8 \cup Dizziness\,With\,Eye\,Open$), resulting in 6 symptoms. Dizziness With Eye Open was found to be the relevant symptom only for test #1. This may be explained by the fact that test #1 is quite different from the other five tests in terms of test design and the duration of the videos. Therefore, we include the symptom Dizziness With Eye Open in the R9 questionnaire.

In order to compare the set of questions with the SSQ, we categorized the symptoms into three categories as described in the original paper introducing the SSQ [18] for all the questionnaires, as shown in Table 5. Here, we are assuming that the three clusters from Kennedy et al. [18] are still valid. For the comparison, we calculate the Pearson Correlation (PCC), so as to identify which of the questionnaires (out of 11) has the maximum correlation with the SSQ. Therein, the N, O, and D analysis only serves as an additional test whether the right symptoms were left out. We used the same scoring method as described in the original SSQ [18], see Eq. 1. From our tests, we derived nine different questions and two questionnaires (VRSQ and CSQ) are from the state-of-the-art. We conduct PCC analysis between the simplified scoring and the original scoring for all the 11 questionnaires for all the three sub-categories of SSQ (N, O and D) and the Total Score respectively on each of the tests. Table 6 shows Pearson Correlation Coefficient for all the tests. Most of the questionnaires have high PCC values for each of the tests except for test #2, where all questionnaires have relatively lower PCC values. R4 and VRSQ have a negative correlation with the Nausea component of the SSQ for the test #2. With the increase in the session duration, the SSQ scores increases. But with R4 and VRSQ, the scores for Nausea component decreases with time due to usage of a relatively less number of symptoms as compared to the original SSQ.

Table 7 shows the average PCC values over all tests and categories of the SSQ (N, O and D) and TS for each of the questionnaires. It can be clearly seen that R7 (9 symptoms) has the highest PCC with the SSQ as compared to the other questionnaires for N, O and TS. Other questionnaires such as R2, R5 and R6 have higher correlations with N, O, D and TS and contain only 6 symptoms. R9 also has a higher correlation and contains 6 symptoms as well. VRSQ and CSQ have the lowest PCC as compared to the rest of the questionnaires for Nausea category and the average PCC over N, O, D and TS. The possible reason could be as VRSQ and CSQ are derived using only VR contents not 360° videos and both of the questionnaires have used only a very limited VR dataset.

Based on the analysis, we recommend using the R7 simplified SSQ as it has the highest PCC as compared to the state-of-the-art questionnaires VRSQ and CSQ, and the number of symptoms (9) remains the same. Furthermore, we recommend using R9 simplified SSQ as it has a lower number of symptoms

Table 4: Obtained magnitude for each of the symptoms after applying PCA on the raw scores of Simulator Sickness for each Subject for Test # 1

Subject ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	Percentage
No. of PCs	5	5	4	5	4	3	6	3	2	5	N/A	3	4	4	3	3	2	3	1	5	4	3	3	3	3	4	N/A	4	
General Discomfort	0.697	0.61	0.529	0.286	0.512	0.766	0.824	0.512	0	0.742	0	0.537	0.531	0.91	0.267	0	0.396	0	0	0.287	0.52	0.489	0.337	0.675	0.305	0.491	0	0.523	57
Fatigue	0.871	0.454	0.981	0.83	0.233	0	0	0.493	0	0.817	0	0.116	0.407	0.861	0	0	0	0.886	0	0.747	0	0.165	0.188	0.31	0	0.666	0	0.468	43
Headache	0.924	0.066	0.321	0.253	0.547	0	0.708	0.399	0	0	0	0	0	0.175	0.182	0	0	0.44	0	0.431	0.287	0	0.207	0.198	0	0.433	0	0.497	25
Eye Strain	0.952	0.29	0.743	0.869	0.24	0.1	0.896	0.166	0.61	0.746	0	0	0	0.398	0.835	0.369	0.918	0	0	0.91	0.51	0.437	0.302	0.266	0.545	0.963	0	0.613	50
Difficulty Focusing	0	0.518	0.694	0.553	0.75	0.978	0.59	0.237	0.937	0.592	0	0.406	0.559	0	0.657	0.989	0	0	0	0.761	0.477	0.437	0.564	0.653	0.894	0.697	0	0.62	71
Salivation Increasing	0	0.91	0.321	0.296	0	0	0.869	0.434	0	0	0	0.565	0	0.453	0	0	0	0	0	0	0.603	0.622	0	0	0	0.333	0	0.289	25
Sweating	0	0	0	0	0	0	0	0.206	0	0	0	0	0	0.154	0	0	0	0	0	0	0.556	0	0	0.213	0	0	0	0.308	4
Nausea	0.484	0.112	0	0	0	0.21	0.22	0.402	0	0	0	0	0.632	0.159	0.505	0	0	0	0	0.694	0.725	0.164	0.624	0.177	0.498	0.577	0	0.555	36
Difficulty Concentrating	0.64	0.827	0	0.83	0.6	0.473	0.594	0.314	0	0.738	0	0	0.646	0.68	0.3	0.956	0	0.245	0	0.511	0.441	0.489	0.442	0.574	0.175	0.486	0	0.639	61
Fullness of Head	0.89	0.849	0	0.931	0.248	0	0.213	0.348	0	0.62	0	0.537	0	0.551	0	0	0	0	0	0.966	0.371	0.501	0.371	0.394	0	0.135	0	0.587	32
Blurred Vision	0.446	0.892	0.908	0.825	0.796	0.981	0.855	0.32	0.866	0.832	0	0.537	0.866	0	0.887	0.977	0.396	0.998	0	0.607	0.746	0.987	0.668	0.949	0.997	0.404	0	0.567	79
Dizziness with Eye Open	0.534	0.632	0.833	0.505	0	0.465	0.259	0.353	0	0.394	0	0.994	0.702	0.741	0.636	0	0.918	0.695	1	0.508	0.647	0.26	0.489	0.383	0.644	0.613	0	0.474	64
Dizziness with Eye Closed	0.195	0.686	0	0.66	0	0	0	0	0	0.584	0	0	0.718	0.442	0	0	0	0.695	0	0.508	0.315	0.399	0.489	0	0	0	0	0.552	32
Vertigo	0.195	0.372	0	0.296	0.944	0	0.274	0.325	0	0.87	0	0.806	0.765	0.49	0.478	0.131	0	0	0	0.452	0.677	0.489	0.667	0.442	0.305	0.692	0	0.343	43
Stomach Awareness	0	0.126	0	0	0	0	0.944	0.73	0	0	0	0	0.283	0.486	0	0	0	0	0	0	0.31	0.103	0.391	0.326	0.175	0	0	0.53	14
Burping	0	0	0	0	0.904	0	0.836	0.879	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.391	0	0	0	0	0.111	11

Table 5: Different Simplified Simulator Sickness Questionnaires Obtained from Several Tests Using PCA.

Symptoms	R1 (Test 1)			R2 (Test 2)			R3 (Test 3)			R4 (Test 4)			R5 (Test 5)			R6 (Test 6)			R7 (Union)			R8 (Intersection1)			R9 (Intersection2)			VRSQ			CSQ		
	N	O	D	N	O	D	N	O	D	N	O	D	N	O	D	N	O	D	N	O	D	N	O	D	N	O	D	N	O	D	N	O	D
General Discomfort	×	×		×	×					×	×		×	×		×	×		×	×		×	×		×	×		×	×				
Fatigue					×			×			×			×			×			×			×			×			×				
Headache											×									×									×				
Eye Strain		×			×			×			×			×			×			×			×			×			×			×	
Difficulty Focusing		×	×		×	×								×	×		×	×		×	×								×	×		×	×
Salivation Increasing																																		
Sweating																																		
Nausea																															×		×
Difficulty Concentrating	×	×		×	×		×	×					×	×		×	×		×	×		×	×		×	×								
Fullness of Head						×			×			×			×			×			×			×			×			×			×	×
Blurred Vision		×	×																	×	×			×					×	×		×	×
Dizziness with Eye Open			×																	×							×						×
Dizziness with Eye Closed																														×			×	
Vertigo																														×			×	
Stomach Awareness																																		
Burping																																		

Table 6: Pearson Correlation Coefficient between N, O, D and TS of the SSQ and several reduced questionnaires obtained from PCA.

	Test 1				Test 2				Test 3				Test 4				Test 5				Test 6				Test 7			
Questionnaires	N	O	D	TS	N	O	D	TS	N	O	D	TS	N	O	D	TS	N	O	D	TS	N	O	D	TS	N	O	D	TS
R1	0.969	0.991	0.956	0.982	0.824	0.873	0.646	0.817	0.98	0.998	0.977	1	0.969	0.992	0.991	0.998	0.976	0.991	0.878	0.977	0.973	0.997	0.985	0.999	0.957	0.988	0.985	0.977
R2	0.969	0.991	0.925	0.988	0.824	0.999	0.685	0.977	0.98	0.999	0.941	0.994	0.969	0.99	0.98	0.996	0.976	0.998	0.972	0.994	0.973	0.991	0.985	0.984	0.957	0.997	0.954	0.988
R3	0.875	0.953	0.865	0.941	0.873	0.982	0.203	0.981	0.996	0.997	0.968	1	0.979	0.972	0.757	0.982	0.937	0.987	0.85	0.989	0.897	0.933	0.99	0.95	0.953	0.987	0.967	0.998
R4	0.967	0.943	0.865	0.962	-0.746	0.943	0.203	0.911	0.9	0.995	0.968	0.991	0.959	0.998	0.757	0.991	0.903	0.995	0.85	0.982	0.824	0.947	0.99	0.951	0.948	0.992	0.967	0.995
R5	0.969	0.991	0.925	0.988	0.824	0.999	0.685	0.977	0.98	0.999	0.941	0.994	0.969	0.99	0.98	0.996	0.976	0.998	0.972	0.994	0.973	0.991	0.985	0.984	0.957	0.997	0.954	0.988
R6	0.969	0.991	0.925	0.988	0.824	0.999	0.685	0.977	0.98	0.999	0.941	0.994	0.969	0.99	0.98	0.996	0.976	0.998	0.972	0.994	0.973	0.991	0.985	0.984	0.957	0.997	0.954	0.988
R7 (Union)	0.969	1	0.967	0.989	0.824	1	0.938	0.982	0.98	1	0.997	1	0.969	1	0.974	0.992	0.976	1	0.988	0.994	0.973	1	0.989	1	0.957	1	0.985	0.993
R8 (Intersection 1)	0.969	0.958	0.865	0.971	0.824	0.985	0.203	0.988	0.98	0.998	0.968	0.994	0.969	0.991	0.757	0.997	0.976	0.996	0.85	0.999	0.973	0.989	0.99	0.99	0.957	0.99	0.967	0.987
R9 (Intersection 2)	0.969	0.958	0.92	0.972	0.824	0.985	0.428	0.992	0.98	0.998	0.998	0.999	0.969	0.991	0.912	0.995	0.976	0.996	0.993	0.992	0.973	0.989	0.985	0.989	0.957	0.99	0.99	0.99
VRSQ	0.967	0.993	0.984	0.985	-0.746	0.994	0.963	0.991	0.9	0.999	0.996	0.995	0.959	1	0.982	0.988	0.903	0.997	0.996	0.986	0.824	0.997	0.991	0.99	0.948	1	0.979	0.996
CSQ	0.95	0.945	1	0.985	0.461	0.956	1	0.989	0.956	0.996	1	0.999	0.307	0.997	1	0.999	0.372	0.994	1	0.998	0.951	0.985	1	0.959	0.901	0.991	1	0.992

$$
\begin{aligned}
Uneasiness &= \frac{(GeneralDiscomfort + Fatigue + EyeStrain + Headache + DifficultyConcentrating + FullnessOfHead)}{18} \times 100 \\
VisualDiscomfort &= \frac{(DifficultyFocusing + BlurredVision)}{6} \times 100 \\
LossofBalance &= \frac{(DizzinessWithEyeOpen)}{3} \times 100 \\
TotalScore &= \frac{(Uneasiness + VisualDiscomfort + LossOfBalance)}{3}
\end{aligned}
\tag{3}
$$

$$
\begin{aligned}
Uneasiness &= \frac{(Fatigue + EyeStrain + DifficultyConcentrating + FullnessOfHead)}{12} \times 100 \\
LossofBalance &= \frac{(GeneralDiscomfort + DizzinessWithEyeOpen)}{6} \times 100 \\
TotalScore &= \frac{(Uneasiness + LossOfBalance)}{2}
\end{aligned}
\tag{4}
$$

i.e. 6 and still has a higher PCC value. The reason for using R9 instead of R1 or R2 or R5 is it contains the relevant symptoms such as Fatigue and Dizziness with Eye Open. Fatigue is relevant when a person is in a VR session for a longer duration of time. Whereas Dizziness with Eye Open is relevant for a subjective test where contents are being used that contain high amounts of motion.

In order to find if we have not left any frequently occurred or relevant symptom during PCA, we averaged the raw scores over all subjects and sessions/videos to obtain the average value of a symptom for each test. Figure 1 shows the average value of each symptom for each test. It is interesting to note that the relevant symptoms obtained after applying PCA for each test are contained within the top 8 symptoms having the highest magnitude. Some of the symptoms such as Nausea, Burping, Sweating, Salivation Increasing, Stomach Awareness, Vertigo and Dizziness with Eyes closed were never selected as a relevant symptom, also lie within the bottom 8 symptoms. After identifying the most similar questionnaire w.r.t. SSQ, the next step is to investigate, whether these three symptom clusters are still valid for 360° related studies.

Table 7: Pearson Correlation Coefficient averaged over all tests for different categories of SSQ for several reduced questionnaires obtained using PCA and for VRSQ and CSQ.

	Correlation Value				
Questionnaires	N	O	D	TS	Average_All
R1	0.949	0.974	0.906	0.962	0.947
R2	0.949	0.995	0.915	0.989	0.962
R3	0.926	0.971	0.772	0.974	0.911
R4	0.634	0.97	0.772	0.965	0.835
R5	0.949	0.995	0.915	0.989	0.962
R6	0.949	0.995	0.915	0.989	0.962
R7	**0.949**	**1**	0.975	**0.993**	**0.979**
R8	0.949	0.986	0.772	0.99	0.924
R9	**0.949**	0.986	0.872	0.99	**0.949**
VRSQ	0.634	0.996	0.985	0.989	0.901
CSQ	0.666	0.979	**1**	0.988	0.908

4.2 Exploratory Factor Analysis

As explained in Sec. 1, due to the differences between VR and simulator usage, it is of relevance to analyse whether the underlying clusters or latent factors initially described by Kennedy et al. [18] are applicable to simulator sickness in the context of 360° video. To this aim, we conducted exploratory factor analysis to the SSQ scores obtained across the different tests. In order to apply factor analysis, we average the raw simulator sickness scores across each subject for each symptom and PVS/session. These averaged out values for each test and the corresponding session/PVS are stacked in a single file.

The Kaiser-Meyer-Olkin (KMO) test was performed to find if the data used for factor analysis is suitable or not. In our study, the KMO value found is 0.82, which means that the data can be considered suitable for factor analysis [10, 14].

Exploratory Factor Analysis (EFA) was first performed on the R7 (Union) and R9 (Intersection 2) questionnaires, including varimax rotation. To find out how many factors are sufficient, we perform the hypothesis test that suggests three and two factors are sufficient for the case of R7 and R9, respectively. The variance explained by the three factors (R7) and two components (R9) is 85.4% and 86.1% respectively of the total variance. Tables 8 and 9 show the factor loadings for the varimax orthogonal rotation for individual symptoms for R7 and R9 questionnaires. For the R7 questionnaire, factors with less than 0.53 factor loadings were removed, resulting in a simple structure where each symptom belongs to one factor only, which eases the naming of the factors. Factor 1 loads mainly on aspects of uneasiness such as General Discomfort, Fatigue, Eye Strain, Headache, Difficulty Concentrating and Fullness of head. Factor 2 loads mainly on visual-discomfort-related symptoms such as Difficulty Focusing and Blurred Vision. Factor 2 could also be referred to as visual quality factors, related with both the quality-rating task present in all six tests as well as the types of degradations included in many of these (different resolutions, encoding bitrates, HMDs). Factor 3 loads only on loss of balance and Dizziness with eyes open. One could summarize the factor labels for R7 as Factor 1: Uneasiness, Factor 2: Visual Discomfort and Factor 3: Loss of Balance. A simple average method was used based on [20] to calculate Uneasiness, Visual Discomfort, Loss of Balance and total score as shown in Eq. 3 for the R7 questionnaire.

For the R9 questionnaire, factors with less than 0.60 factor loadings were removed from each component, resulting in a simple structure where each symptom belongs to one factor only, easing the naming of the factors. Factor 1 loads mainly on aspects of uneasiness such as Fatigue, Eye Strain, Difficulty Concentrating and Fullness of head. Factor 2 loads only on loss of balance, General Discomfort and Dizziness with eyes open. One could summarize the factor labels for R9 as Factor 1: Uneasiness and Factor 2: Loss of Balance. Uneasiness, Loss of Balance and total score are calculated based on a simple average method described in [20] as shown in Eq. 4 for the R9 questionnaire.

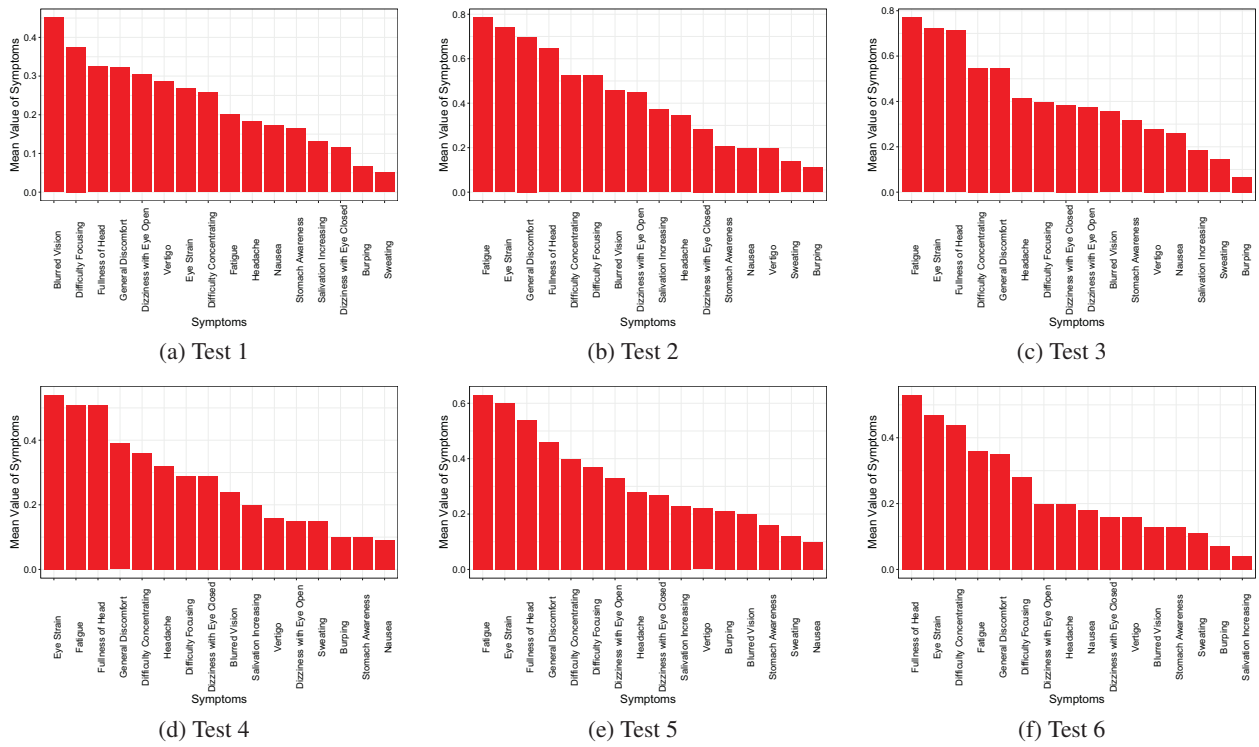
We measured Cronbach's Alpha (α) which is a measure of the internal consistency. For our R7, Cronbach's Alpha is 0.91 and for R9 it was found to be 0.93, which signifies that the questionnaires examined here show sufficient reliability.

Thus, the outcomes from our study do not negate the factor structure from the study conducted by Kenndy et al. [18], especially for the R7 questionnaire. Both of the studies point to the same factors, namely loss of balance (Disorientation), visual discomfort (Oculomotor) and uneasiness (Nausea), contributing to the assessment of simulator sickness, but contains different symptom clusters.

Further, EFA was performed on the original SSQ using our data in order to check if the same set of symptom clusters can be obtained for our tests as those described for the original SSQ [18]. Table 10 shows the factor loadings for the varimax orthogonal rotation for individual symptoms for the SSQ questionnaires. The variance explained by the three factors is 73.5% of the total variance. To find which of the symptom belongs to which cluster, factors with less than 0.30 factor loadings were removed from each factor. Here, we kept the same threshold as defined in the SSQ [18]. It can be seen that the resulting symptom clusters are different from the ones defined by Kennedy et al. for N, O and D. For our case of 360° video, a different set of symptoms were found, belonging to different factors.

4.3 Discussion

The subjective tests (#1 to #6) that we have considered for developing the reduced-set questionnaire were originally targeting research on video quality. From Fig. 1, it can clearly be seen that the strongest effect on participants is that of Blurred vision, Difficulty Focusing, Eye Strain, Difficulty Concentrating and General Discomfort. Due to the variation in resolution and bitrate that have been chosen for test stimulus generation, the SSQ itself may address certain aspects of simulator sickness which are not necessarily simulator sickness issues, but



(a) Test 1 (b) Test 2 (c) Test 3

(d) Test 4 (e) Test 5 (f) Test 6

Figure 1: Mean Value of all Symptoms for several tests.

Table 8: Varimax Factors For Reduced SSQ (Union)

Symptom	Factor 1	Factor 2	Factor 3
General Discomfort	**0.701**	0.385	0.52
Fatigue	**0.953**	0	0.148
Headache	**0.826**	0.116	0.298
Eye Strain	**0.938**	0.136	0
Difficulty Focusing	0.341	**0.923**	0.166
Difficulty Concentrating	**0.786**	0.222	0.343
Fullness of Head	**0.897**	0	0.224
Blurred Vision	-0.101	**0.801**	0.202
Dizziness with Eye Open	0.298	0.404	**0.766**

Table 9: Varimax Factors For Reduced SSQ (Intersection 2)

Symptom	Factor 1	Factor 2
General Discomfort	0.595	**0.747**
Fatigue	**0.916**	0.269
Eye Strain	**0.913**	0.223
Difficulty Concentrating	**0.720**	0.518
Fullness of Head	**0.867**	0.317
Dizziness with Eye Open	0.176	**0.887**

may be due to the quality-related characteristics investigated in our studies. Then, these quality-related issues may translate into simulator sickness issues, when rating corresponding symptoms in the simulator sickness questionnaire. To further analyze this hypothesis, we used the data from a further test (Test #7), which is publicly available. That test focused only on collecting viewing behavior from subjects when watching 360° videos on an HMD, collecting simulator sickness data for complementary analysis. In that test, no variation in resolution and bitrate were applied [8]. Inspection of this additional test revealed that users still reported high scores for Blurred vision, Difficulty Focusing, Eye Strain, Difficulty Concentrating and General Discomfort. Therefore, it can be assumed that the quality-related issues introduced in the tests #1 to #6 may not have been translated into simulator sickness issues, when rating corresponding symptoms in the simulator sickness questionnaire. Furthermore, we compare all the reduced questionnaires (R1 to R9), VRSQ and CSQ using PCC analysis including Test #7. It can be seen from Table 6 that all the questionnaires including VRSQ and CSQ show high correlations with N, O, D and TS. This observation supports the conclusion that the SSQ or a corresponding simplified questionnaire can be instantiated in the same way both for 360° video tests including dedicated media-quality-related questions and for tests without focus on visual quality. Also, the proposed R7 questionnaire can be used in the contexts that use 360° videos and 360° audiovisual stimuli, as R7 questionnaire is designed using all tests (#1 to #6) and contains all of the symptoms from test 1 which used audiovisual stimuli. Hence, for the range of conditions we have applied we consider it useful.

We are aware that there are some limitations of the simplified questionnaire. It is noted that the results in this paper are obtained by simulating the usage of the reduced-set questionnaires, simply omitting the questions that are proposed to be left out for 360° video testing. The proposed 9-attribute questionnaire will need to be verified in concrete tests with the

Table 10: Varimax Factors For SSQ (Original)

Symptom	Factor 1 (N)	Factor 2 (D)	Factor 3 (O)
General Discomfort	**0.68**	**0.583**	**0.303**
Fatigue	**0.968**	0.112	0
Headache	**0.819**	**0.349**	0
Eye Strain	**0.918**	0	0.205
Difficulty Focusing	0.268	**0.403**	**0.872**
Salivation Increasing	**0.619**	**0.317**	**0.317**
Sweating	**0.518**	0.151	0
Nausea	0.2	**0.755**	0.28
Difficulty Concentrating	**0.759**	**0.426**	0.178
Fullness of Head	**0.895**	0.22	0
Blurred Vision	-0.138	**0.347**	**0.728**
Dizziness with Eye Open	0.299	**0.802**	0.228
Dizziness with Eye Closed	**0.8**	**0.361**	0
Vertigo	0	**0.93**	0.25
Stomach Awareness	**0.442**	**0.67**	0.184
Burping	**0.444**	0	0

reduced set of questions in the future[6]. Also, the simplified questionnaire was developed using a specific set of 360° videos, and the duration of an individual test sessions were short.

5 Conclusions

In this study, we attempt to develop and validate efficient questionnaires for the assessment of simulator sickness in the context of 360° videos. To this aim, we analyzed the data of previously published six studies conducted by our group. This consists of results obtained from more than 150 participants covering 360° videos with wide ranging spatio-temporal complexity tested with different HMDs with different test designs and paradigms. From this, we derived nine different reduced questionnaires from the original SSQ by applying PCA (per-subject) on the results obtained from the six different subjective tests. We conducted a PCC analysis to check the applicability of the proposed reduced questionnaires in comparison to the SSQ, and how they compare with the CSQ [11] and VRSQ [20] proposed in literature. Experimental results suggest that the proposed R7 questionnaire shows the highest PCC with the SSQ in terms of the initial clusters. Further, results of factor analysis showed that nine symptom-related attributes appear to be sufficient to represent the three dimensions of SSQ, as some symptoms included in the SSQ are not relevant for 360° video studies and can be expected to capture all the relevant symptoms just as the SSQ does and hence can be used as an alternative to the SSQ for studies related to 360° videos. Therefore, based on the analysis, we propose using the simplified SSQ as it contains a lower number of questions (9 instead of 16), with a higher correlation with the original SSQ as compared to VRSQ and CSQ. To conclude, results from our study imply that the simplified version of the SSQ focuses on the symptoms that are relevant for 360° video related studies, and can be used efficiently, enabling larger-scale testing.

As future work, we propose to further evaluate the SSQ with a broader range of contents such as VR and CGI-based 360° videos and also construct an even more holistic yet sufficient subjective measurement tool. Furthermore, we will extend the duration of the VR session to enhance the validation of the different questionnaires.

Acknowledgment

This work has partially been supported by the CYTEMEX project funded by the Free State of Thuringia, Germany (FKZ: 2018-FGI-0019).

[6]This is due to the current pandemic and associated lockdown of our institution, which has limited the ability to run dedicated additional laboratory tests.

References

[1] A. Abbas and B. Adsumilli. *AHG8: New GoPro Test Sequences for Virtual Reality Video Coding*. JVET-D0026, 4th Meeting: Chengdu, China, 15-21 October, 2016.

[2] E. Asbun, Y. He, Y. He, and Y. Ye. *AHG8: InterDigital Test Sequences for Virtual Reality Video Coding*. JVET-D0039, 4th Meeting: Chengdu, China, 15-21 October, 2016.

[3] P. Bimberg, T. Weissker, and A. Kulik. On the usage of the simulator sickness questionnaire for virtual reality research. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, March 2020.

[4] S. Bouchard, G. Robillard, and P. Renaud. *Revising the factor structure of the simulator sickness questionnaire*. Summer. Annu. Rev. Cybertherapy Telemed. 5, 128137, 2007.

[5] S. Bruck and P. A. Watters. The factor structure of cybersickness. *Displays*, 32(4):153 – 158, 2011. Visual Image Safety. doi: 10.1016/j.displa.2011.07.002

[6] K. Carnegie and T. Rhee. Reducing visual discomfort with hmds using dynamic depth of field. *IEEE Computer Graphics and Applications*, 35(5):34–41, Sep. 2015. doi: 10.1109/MCG.2015.98

[7] M. S. Dennison, A. Z. Wisti, and M. DZmura. Use of physiological signals to predict cybersickness. *Displays*, 44:42 – 52, 2016. Contains Special Issue Articles Proceedings of the 4th Symposium on Liquid Crystal Photonics (SLCP 2015). doi: 10.1016/j.displa.2016.07.002

[8] S. Fremerey, A. Singla, K. Meseberg, and A. Raake. Avtrack360: An open dataset and software recording people's head rotations watching 360-degree; videos on an hmd. In *Proceedings of the 9th ACM MMSys*, 2018. doi: 10.1145/3204949.3208134

[9] P. J. Gianaros, E. R. Muth, J. T. Mordkoff, M. E. Levine, and R. M. Stern. A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation, space, and environmental medicine*, 72 2:115–9, 2001.

[10] K. HF. *A Second-Generation Little Jiffy*. Psychometrika;35(4):401-15, 1970.

[11] W. B. S. Iii. *Psychometric Evaluation of the Simulator Sickness Questionnaire as a Measure of Cybersickness*. Doctoral dissertation, Iowa State University, 2017.

[12] ITU-R. *Recommendation ITU-R BT.500-13: Methodology for the subjective assessment of the quality of television pictures*. January 2012.

[13] ITU-T. *Recommendation ITU-T P.910: Subjective video quality assessment methods for multimedia applications*. April 2008.

[14] H. Kaiser, J. Little, and I. Mark. *Educational and Psychological Measurement*. 34:111-7, 1974.

[15] S. Katsigiannis, R. Willis, and N. Ramzan. A qoe and simulator sickness evaluation of a smart-exercise-bike virtual reality system via user feedback and physiological signals. *IEEE Transactions on Consumer Electronics*, 65(1):119–127, 2019.

[16] K. Keerthi Vasan and B. Surendiran. Dimensionality reduction using principal component analysis for network intrusion detection. *Perspectives in Science*, 8:510 – 512, 2016. Recent Trends in Engineering and Material Sciences. doi: 10.1016/j.pisc.2016.05.010

[17] R. Kellogg, R. Kennedy, and A. Graybiel. Motion sickness symptomatology of labyrinthine defective and normal subjects during zero gravity maneuvers. *Aerospace medicine*, 36:315318, April 1965.

[18] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation*

Psychology, 3(3):203–220, 1993.

[19] B. Keshavarz and H. Hecht. Validating an efficient method to quantify motion sickness. *Human Factors*, 53(4):415–426, 2011. PMID: 21901938. doi: 10.1177/0018720811403736

[20] H. K. Kim, J. Park, Y. Choi, and M. Choe. Virtual reality sickness questionnaire (vrsq): Motion sickness measurement index in a virtual reality environment. *Applied Ergonomics*, 69:66 – 73, 2018. doi: 10.1016/j.apergo.2017.12.016

[21] E. M. Kolasinski, S. A. Jones, R. S. Kennedy, and R. D. Gilson. Postural stability and its relation to simulator sickness. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 38, pp. 980–980. SAGE Publications Sage CA: Los Angeles, CA, 1994.

[22] A. Raake, A. Singla, R. R. R. Rao, W. Robitza, and F. Hofmeyer. Sisimo: Towards simulator sickness modeling for 360 videos viewed with an hmd. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 583–584, 2020.

[23] L. Rebenitsch and C. Owen. Review on cybersickness in applications and visual displays. *Virtual Real.*, 20(2):101125, June 2016. doi: 10.1007/s10055-016-0285-9

[24] N. Salem and S. Hussein. Data dimensional reduction and principal components analysis. *Procedia Computer Science*, 163:292 – 299, 2019. 16th Learning and Technology Conference 2019Artificial Intelligence and Machine Learning: Embedding the Intelligence. doi: 10.1016/j.procs.2019.12.111

[25] S. Schwarz, A. Aminlou, and et al. *Tampere pole vaulting sequence for virtual reality video coding*. JVET-D0143, 4th Meeting: Chengdu, China, 15-21 October, 2016.

[26] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson. Virtual reality induced symptoms and effects (vrise): Comparison of head mounted display (hmd), desktop and projection display systems. *Displays*, 29(2):58 – 69, 2008. Health and Safety Aspects of Visual Displays. doi: 10.1016/j.displa.2007.09.005

[27] A. Singla, S. Fremerey, W. Robitza, P. Lebreton, and A. Raake. Comparison of subjective quality evaluation for hevc encoded omnidirectional videos at different bit-rates for uhd and fhd resolution. In *Proceedings of the on Thematic Workshops of ACM Multimedia*, 2017. doi: 10.1145/3126686.3126768

[28] A. Singla, S. Fremerey, W. Robitza, and A. Raake. Measuring and comparing qoe and simulator sickness of omnidirectional videos in different head mounted displays. In *9th International Conference on QoMEX*, pp. 1–6, May 2017. doi: 10.1109/QoMEX.2017.7965658

[29] A. Singla, R. R. R. Rao, S. Gring, and A. Raake. Assessing media qoe, simulator sickness and presence for omnidirectional videos with different test protocols. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1163–1164, March 2019. doi: 10.1109/VR.2019.8798291

[30] A. Singla, W. Robitza, and A. Raake. Comparison of subjective quality evaluation methods for omnidirectional videos with dsis and modified acr. *Electronic Imaging*, 2018:1–6, 2018. doi: 10.2352/ISSN.2470-1173.2018.14.HVEI-525

[31] A. Singla, W. Robitza, and A. Raake. Comparison of subjective quality test methods for omnidirectional video quality evaluation. In *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*, pp. 1–6, Sept 2019.

[32] K. M. Stanney, R. S. Kennedy, and J. M. Drexler. Cybersickness is not simulator sickness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 41(2):1138–1142, 1997. doi: 10.1177/107118139704100292

[33] T. A. Stoffregen and L. J. Smart. Postural instability precedes motion sickness. *Brain Research Bulletin*, 47:437–448, 1998.

[34] W. Sun and R. Guo. *Test Sequences for Virtual Reality Video Coding from LetinVR*. JVET-D0179, 4th Meeting: Chengdu, China, 15-21 October, 2016.

[35] T. Tan, M. Mrak, R. Weerakkody, N. Ramzan, V. Baroncini, G. Sullivan, J.-R. Ohm, and K. McCann. *HEVC subjective video quality test results*. Institution of Engineering and Technology, United Kingdom, 2014. Location: Amsterdam, Netherlands Conference date: 11-15 Sept. 2014. doi: 10.1049/ib.2014.0029

[36] H. T. T. Tran, N. P. Ngoc, C. T. Pham, Y. J. Jung, and T. C. Thang. A subjective study on qoe of 360 video for vr communication. *IEEE 19th International Workshop on MMSP*, pp. 1–6, 2017.

[37] S. Weech, S. Kenny, and M. Barnett-Cowan. Presence and cybersickness in virtual reality are negatively related: a review. *Frontiers in psychology*, 10:158, 2019.

[38] A. H. Wertheim, J. E. Bos, and A. J. Krul. *Predicting motion induced vomiting from subjective misery (MISC) ratings obtained in 12 experimental studies*. Report TNO-TM-01-A066, Soesterbery, NL: TNO Human Factors Research Institute., 2001.