



Methodology to Assess Quality, Presence, Empathy, Attitude, and Attention in Social VR: International Experiences Use Case

Marta Orduna, Pablo Pérez, Jesús Gutiérrez, Narciso García

moc@gti.ssr.upm.es

- VR technology is achieving great interest in applications for **social purposes**



It allows the **transmission of non-verbal signals** such as facial expressions or body postures exchanged during a conversation that influences the effectiveness of face-to-face communications

So, we need **methodologies** to jointly assess **technical** parameters such as video quality and **socioemotional** features such as presence

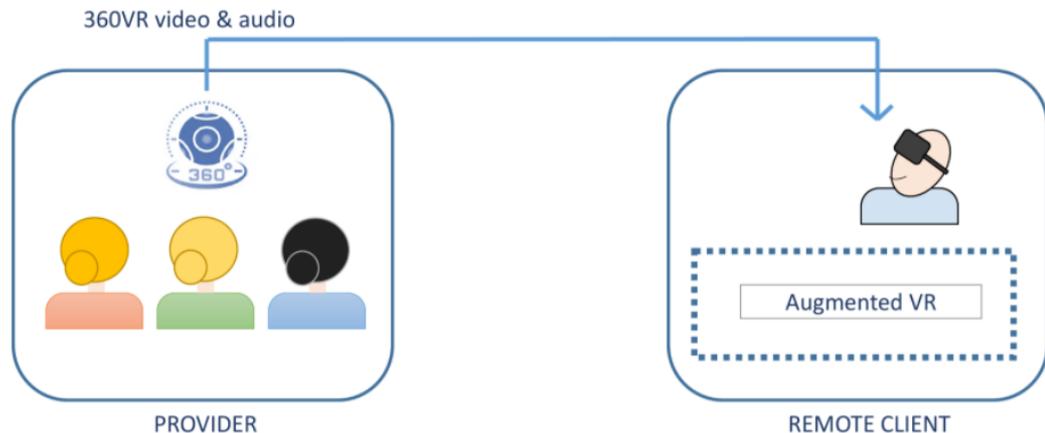


RQ1: Is it possible to evaluate video quality in videos of long duration designed for the evaluation of socioemotional features?

RQ2: Which technical aspects, such as the position of the camera, the type of conversation, the video quality, the acquisition perspective, etc., influence socioemotional features?

RQ3: Which interactive elements can be provided to the user to improve some socioemotional aspects such as presence or attention?



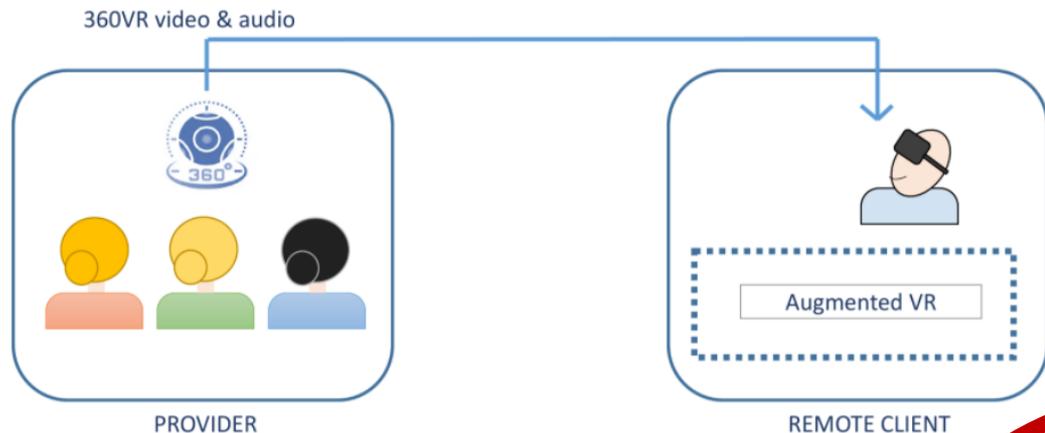


- Observers visualized pre-recorded 360 videos with fluctuations of quality, simulating a VR streaming communication



- Each participant was assigned a condition but all the participants visualized the same PVSs

Condition	Assessment		Interactive element
	Quality	Socioemotional	Hands
A	X	X	
B		X	
C		X	X



- Observers visualized pre-recorded 360 videos with fluctuations of quality, simulating a VR streaming communication



- Each participant was assigned a condition but all the participants visualized the same PVSs



Student Experiences Around the World (SEAW) dataset



Coffee shop



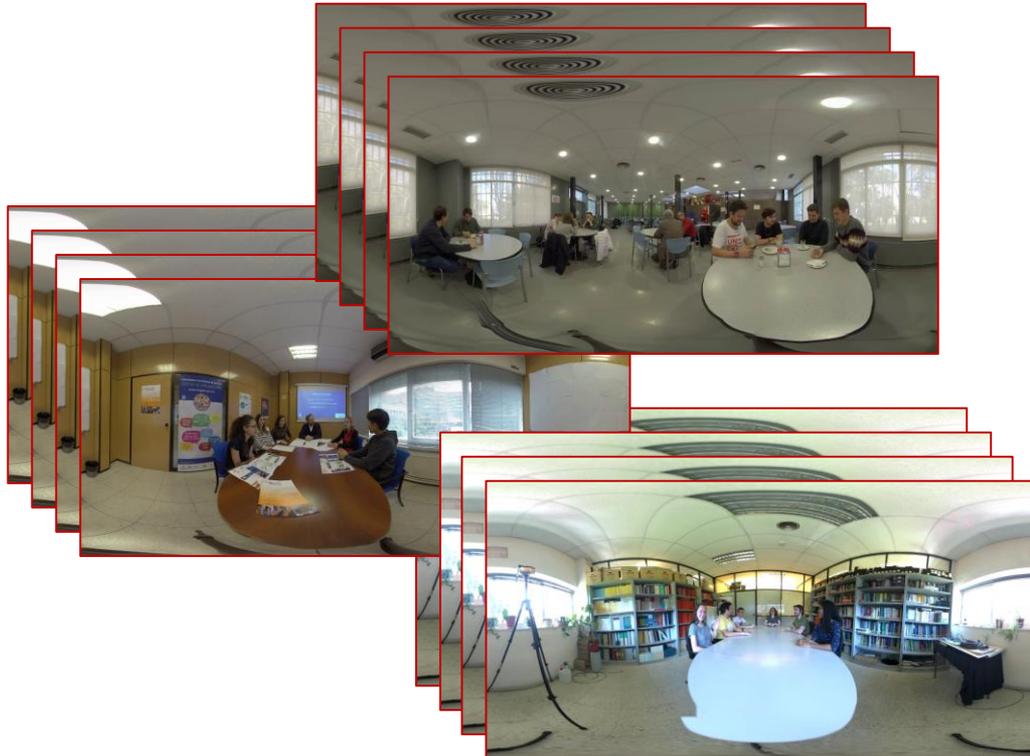
International office



Study in Spain

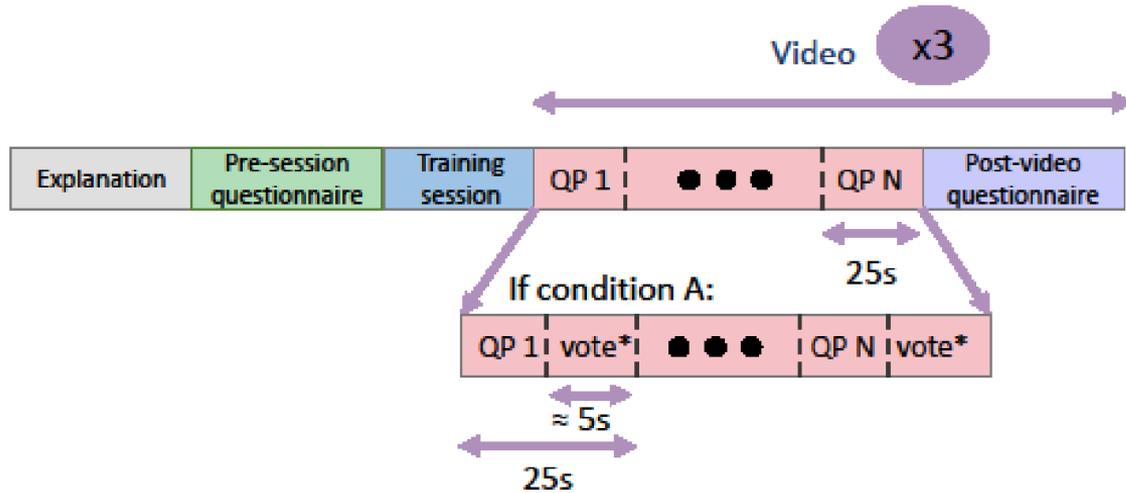
Name	Genre	Perspective-taking	Description
Coffee shop	Everyday conversation	Observer	<i>A coffee conversation between foreign and local students about cultural differences</i>
International office	Educational	Actor	<i>A presentation given by a professor to students about the foreign application process</i>
Study in Spain	Discussion	Actor	<i>A conversation about the differences between transport and rental prices in different countries</i>

QP: 15, 22, 27, 32, 37, and 42



Student Experiences Around the World (SEAW) dataset

- Sequences were encoded with HEVC switching to a different QP each 25 seconds (randomized) to create **one PVS per content**
- Participants from condition A rated the **quality of each one of the 25-seconds** along the whole sequence
- Duration: \approx 5 minutes





Methodology

Condition	Pre-questionnaire (once)			During each content		Post-questionnaire (for each content)			
	Personal information	Empathy (IRI)	Attitude	Quality (SSDQE)	Quality (ACR)	Attention	Attitude	Spatial and Social Presence	Notes
A	X	X	X	X	X	X	X	X	
B	X	X	X		X	X	X	X	
C	X	X	X		X	X	X	X	X

- **Personal information**
- **Empathy:** Interpersonal Reactivity Index (IRI)
- **Attitude:** designed questionnaire based on Facet theory
- **Quality:** SSDQE and ACR
- **Attention:** three questions that had pass/fail answers
- **Spatial and social presence:** aggregate measure of five items
- **Notes:** have your annotations helped you to solve the questions?



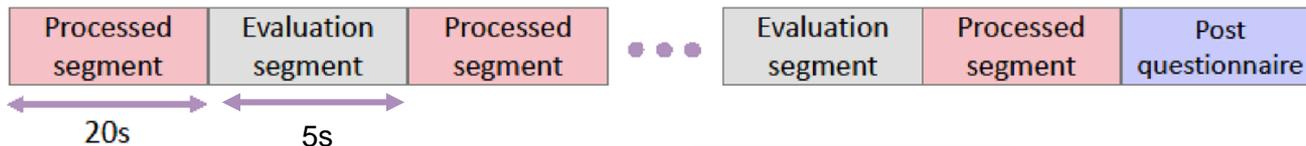
Methodology



Condition	Pre-questionnaire (once)			During each content		Post-questionnaire (for each content)			
	Personal information	Empathy (IRI)	Attitude	Quality (SSDQE)	Quality (ACR)	Attention	Attitude	Spatial and Social Presence	Notes
A	X	X	X	X	X	X	X	X	
B	X	X	X	X	X	X	X	X	
C	X	X	X	X	X	X	X	X	X

SSDQE

Structure of the test sequences used with Single-Stimulus Discrete Quality Evaluation (SSDQE) methodology



ACR

the aggregate quality was asked in the post-questionnaire:

Five Grade Scale - Quality	
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad



Methodology



Condition	Pre-questionnaire (once)			During each content		Post-questionnaire (for each content)			Notes
	Personal information	Empathy (IRI)	Attitude	Quality (SSDQE)	Quality (ACR)	Attention	Attitude	Spatial and Social Presence	
A	X	X	X	X	X	X	X	X	
B	X	X	X		X	X	X	X	
C	X	X	X		X	X	X	X	X

Spatial Presence

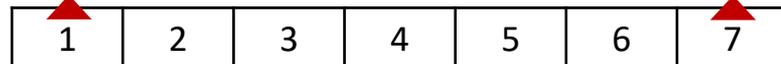
- I felt I was present in the places shown in the video
- I felt surrounded by the actions in the video
- I felt I was sitting by the table at the place of the video
- I felt I could have reached out and touched the items on the table of the video
- I felt that all my senses were stimulated at the same time

Social Presence

- I felt that people were talking to me
- I felt that I was listening to the others in the video
- I felt I was present with the other people in the video
- I felt like the people in the video could see me
- I felt I was actually interacting with other people

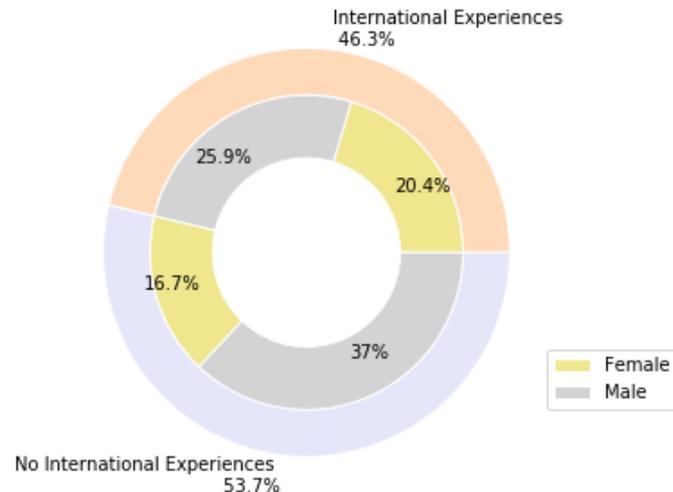
Strongly disagree

Strongly agree



A total of 54 observers:

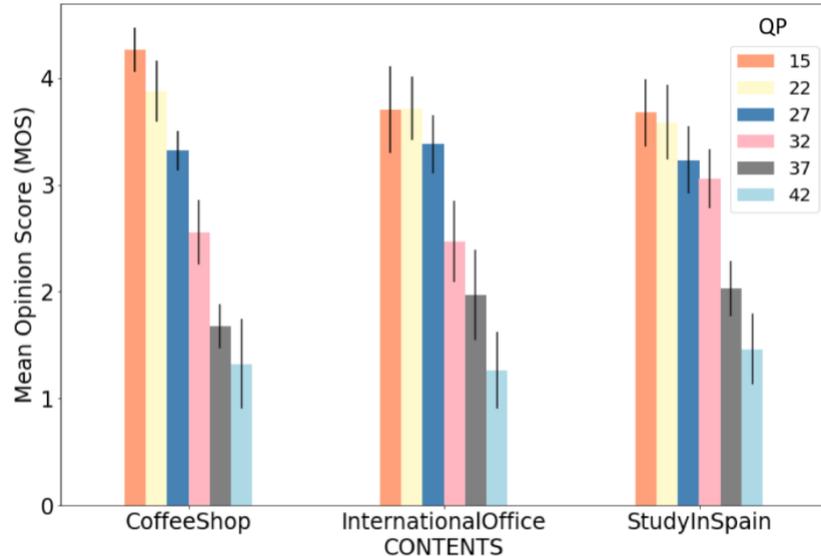
- 20 females
- 34 males
- age range between 17 and 26 years
- **international experiences or nationalities from 15 countries in Europe, America, and Asia**





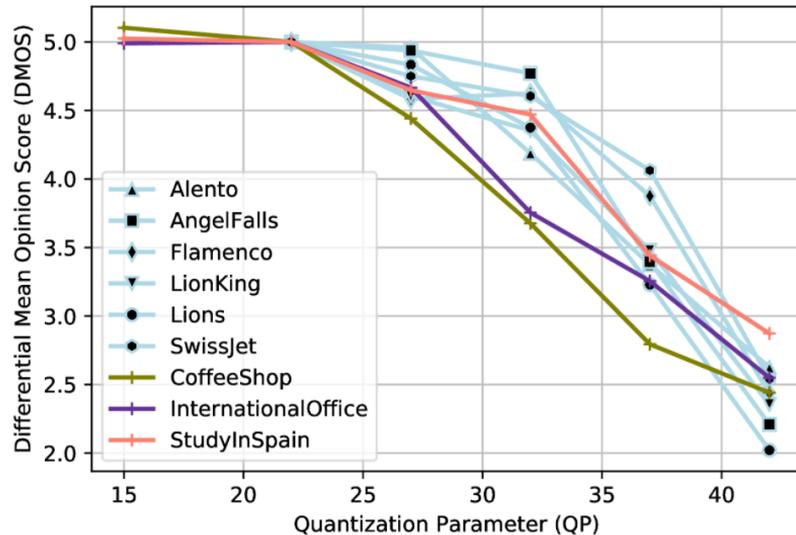
Results

H1: video quality evaluation can be adapted to long-duration videos designed for socioemotional features assessment purposes



- **not significant difference** between QP values of **15 and 22**
- **significant difference** between the **rest of QP values**

H1: video quality evaluation can be adapted to long-duration videos designed for socioemotional features assessment purposes



- 😊 **good distribution** of the ratings and a **consistent decrease of the perceived quality when increasing the QP**, as expected in this type of tests

*DMOS was calculated with QP of 22 as hidden reference



Results



Aggregate quality

Questionnaire items	Coffee shop	International office	Study in Spain	Significance
Aggregate quality (5-level scale)	M = 3.111 (SD = .904)	M = 3.222 (SD = .883)	M = 3.481 (SD = .885)	$F_{2,153} = 2.485, p > .05$

Questionnaire items	Condition A	Condition B	Condition C	Significance
Aggregate quality (5-level scale)	M = 3.537 (SD = .719)	M = 3.111 (SD = 1.022)	M = 3.167 (SD = .885)	$F_{2,153} = 3.687, p < .05$

Condition	Assessment		Interactive element
	Quality	Socioemotional	Hands
A	X	X	
B		X	
C		X	X



Results

H1 - conclusions:

- Subjects are able to assess the video quality of individual QPs and the content does not distract them from this task (assess the video quality)
- Growth of the ecological validity of the quality evaluation compared with traditional methods
- Participants that are focused on the quality evaluation along the sequence, change their perspective about the perceived global quality (aggregate quality)

H2: acquisition perspective, type of the conversation, and experimental condition have influence on: spatial and social presence

Questionnaire items	Condition A	Condition B	Condition C	Significance
Aggregate quality (5-level scale)	M = 3.537 (SD = .719)	M = 3.111 (SD = 1.022)	M = 3.167 (SD = .885)	$F_{2,153} = 3.687, p < .05$
Spatial Presence (7-level scale)	M = 5.463 (SD = 1.019)	M = 5.185 (SD = 1.318)	M = 5.411 (SD = .942)	$\chi^2 = .726, p > .05, df=2$
Social Presence (7-level scale)	M = 5.059 (SD = 1.398)	M = 5.133 (SD = 1.287)	M = 5.144 (SD = 1.271)	$\chi^2 = .09, p > .05, df=2$
Change in attitude (7-scale level)	M = 1.931 (SD = 1.026)	M = 2.514 (SD = 1.024)	M = 2.250 (SD = 1.030)	$F_{2,153} = 4.309, p < .05$
Attention (3-level scale)	M = 1.981 (SD = .765)	M = 1.833 (SD = .885)	M = 1.685 (SD = .748)	$F_{2,153} = 1.839, p > .05$

Condition	Assessment		Interactive element
	Quality	Socioemotional	Hands
A	X	X	
B		X	
C		X	X

H2: acquisition perspective, type of the conversation, and experimental condition have influence on: spatial and social presence

Questionnaire items	Coffee shop	International office	Study in Spain	Significance
Aggregate quality (5-level scale)	M = 3.111 (SD = .904)	M = 3.222 (SD = .883)	M = 3.481 (SD = .885)	$F_{2,153} = 2.485, p > .05$
Spatial Presence (7-level scale)	M = 5.326 (SD = 1.173)	M = 5.200 (SD = 1.137)	M = 5.533 (SD = .991)	$\chi^2 = 4.734, p > .05, df=2$
Social Presence (7-level scale)	M = 4.748 (SD = 1.364)	M = 4.752 (SD = 1.280)	M = 5.837 (SD = .964)	$\chi^2 = 39.166, p < .01, df=2$
Change in attitude (7-scale level)	M = 2.292 (SD = .956)	M = 2.046 (SD = 1.064)	M = 2.356 (SD = 1.111)	$F_{2,153} = 1.352, p > .05$
Attention (3-level scale)	M = 2 (SD = .777)	M = 1.704 (SD = .743)	M = 1.796 (SD = .877)	$F_{2,153} = 1.925, p > .05$



A method to simultaneously assess video quality and socioemotional features

- SSDQE is valid to evaluate individual quality variations
- SSDQE does not affect the evaluation of presence or attention



socioemotional features can be assessed despite having the extra task of continuous video quality evaluation



SSDQE does not reduce the observer immersion, making it a real content-immersive method

Limitations:

- SSDQE **affects** the evaluation of **the aggregate quality** of the sequence
- SSDQE has a small **impact** but statistically significant **on the attitude change** of observers
- The experiment has been carried out in a **specific type of content** and context



Contributions

- **SSDQE methodology.** We propose and validate a methodology to jointly assess video quality and presence, empathy, attitude, and attention in VR communications
 - Extension of the experiment in progress: ACR and SSCQE
- **Video quality assessment in immersive communications**
- **Dataset.** We will make publicly available a **Student Experiences Around the World dataset (SEAW-dataset)** of 3 video sources (stereoscopic raw format) designed and acquired specifically for the purposes of the experiment.
 - Additionally, the questionnaires and the associated rates obtained from a diverse and balanced sample of 54 participants are provided

<https://arxiv.org/abs/2103.02550>

Thank you!

Marta Orduna, moc@gti.ssr.upm.es