VQEG

IMG Work Plan - what's next?

2022-12 – Online

Jesús Gutiérrez (Universidad Politécnica de Madrid, Spain)
Pablo Pérez (Nokia, Spain)

Task based / interactive use cases
The problem

- Imagine you have a bi-directional immersive communication system
  - How do you test it?
    - Evaluate effect of technical factors in QoE (e.g. variations of latency / bitrate / etc.)
    - Compare with other systems / experiments
- ITU-T P.920 - Interactive test methods for audiovisual communications
  - Some tasks proposed to evaluate effect of technical factors:
    - E.g.: one of the subjects shows and describes a plastic building block and the other one is required to reproduce it;
  - Centered on video-conference (05/2000)
- ITU-T P.1320 (P.QXM) - QoE Assessment of eXtended Reality (XR) Meetings
  - Best practices for QoE assessment of telemeetings with extended reality elements
  - Not to the detail of proposing evaluation tasks or methodologies

Task based / interactive use cases
Proposal of joint experiment

- Gather a set of **immersive communication systems**, e.g.
  - Real-time 360 video telepresence
  - Social VR with pointcloud transmission / with avatars
  - AR collaboration
- Create an experiment with covers all basic functionalities
  - Conversation between people
  - Discussion about objects in the immersive space
  - Interaction with (local / remote / virtual) objects in the immersive space
- Run a cross-lab experiment using any available collaboration technology
  - “The same” experiment in completely different setups

Target: Creating the “lego-block” experiment for VR/AR/XR

Task based / interactive use cases

Background

- ITU-T P.1320
  - Factors influencing QoE of XR telemeetings
  - Constituents of QoE in XR telemeetings
  - Test methods targeting XR telemeeting QoE
- *Overview, Taxonomy and Good Practises* (article)
  - Characterize types of systems
  - Identify evaluation tasks

Pablo Pérez, Ester González-Sosa, Jesús Gutiérrez, and Narciso García, *Emerging Immersive Communication Systems: Overview, Taxonomy, and Good Practises for QoE Assessmen*t, Frontiers in Signal Processing, vol. 2, article 917684, pp. 1-22, Jul. 2022

ITU-T P.1320
Factors influencing QoE of XR telemeetings

- Human influence factors (HIFs)
  - Perceptual process, cognitive state, relation between participants…
- Context influence factors (CIFs)
  - Communication environment, use case, …
- System influence factors (SIFs)
  - DoF, representation of users, locomotion, positioning…
  - Rendering
  - Network and compression

P.1320
Constituents of QoE in XR telemeetings

- Simulator sickness
- Immersion
- Presence, co-presence and social presence
- Plausibility
- Ethics of XR use
- Carving out mental space
- Fatigue and cognitive load
- Ability to reach goals

P.1320
Test methods targeting XR telemeeting QoE

- Test design
  - System IF: independent variable
  - Human IF / Context IF: controlled and balanced
  - QoE constitutents: dependent variable
- Define a suitable task: SIF variation ➔ QoE variation
  - ITU-T P.1301: summary of tasks for QoE assessment in telemeetings
  - These tasks have not yet been validated for XR.

- Two main objectives
  - Systematically control an independent variable and observe the effects on QoE
  - Test complete black box systems without exploring individual variables
- Test paradigms
  - Conversation/behavior analysis
  - Post-experience questionnaires
  - Physiological measures

Emerging Immersive Communication Systems:
Overview, Taxonomy, and Good Practises for QoE Assessment

- How many **different systems** are there?
  - Basic features of immersive communication systems
- How many different tasks do people use to evaluate those systems?
- Which system factors (independent variables) are typically evaluated?
- Which QoE elements are typically measured?

Pablo Pérez, Ester González-Sosa, Jesús Gutiérrez, and Narciso García, *Emerging Immersive Communication Systems: Overview, Taxonomy, and Good Practises for QoE Assessmen*t, Frontiers in Signal Processing, vol. 2, article 917684, pp. 1-22, Jul. 2022

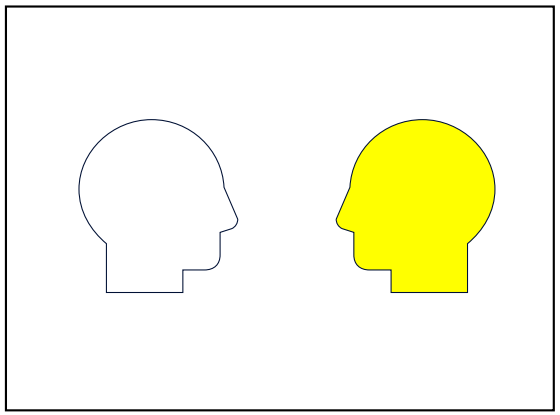
VQEG

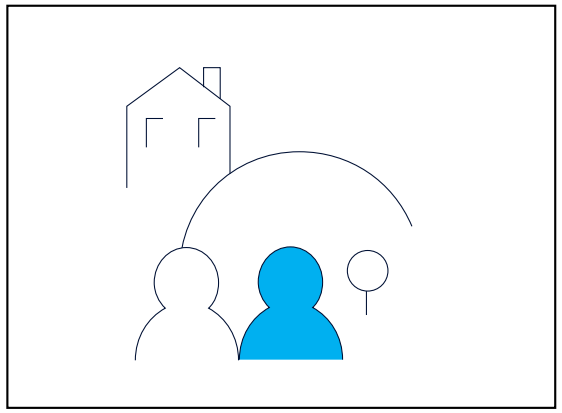
FACE

Visual Communication

Face to face

I see you





VISIT

Remote Presence

Shoulder-to-shoulder

I see what you see

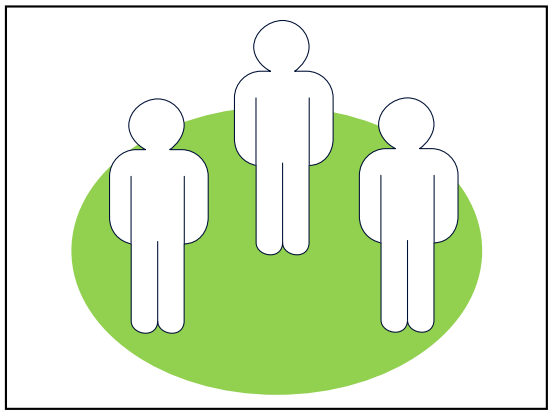
Technological / perceptual building blocks

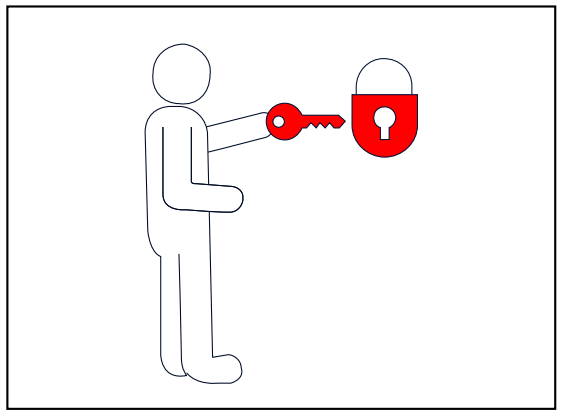
MEET

Shared Immersion

Hand-in-hand

I am with you





MOVE

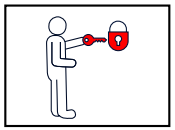
Embodied interaction

Hands-on

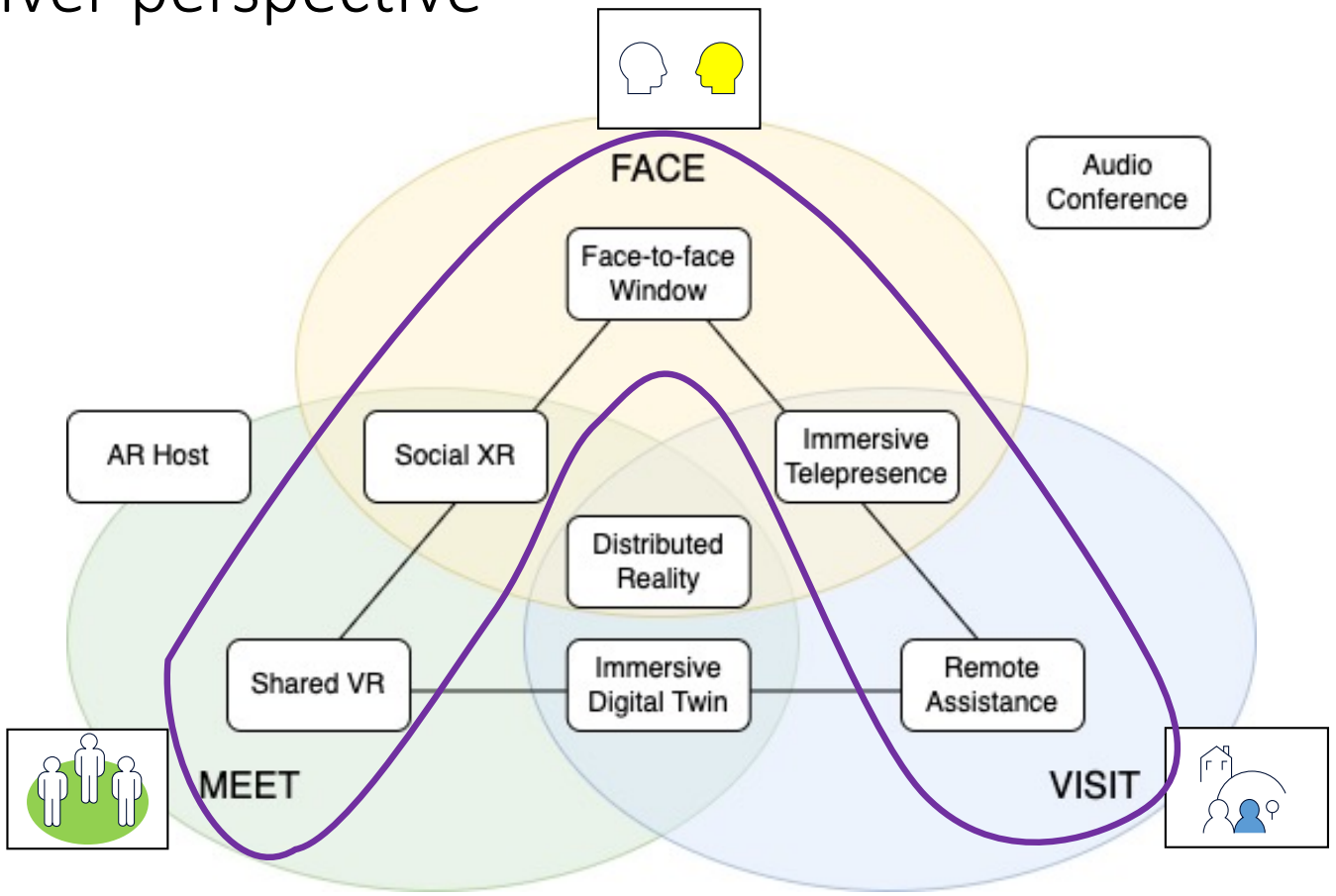
I control objects

Archetypes of XR communication systems

Receiver perspective



MOVE



Archetypes of XR communication systems
Implementation

- Systems of the same archetype have similar technical implementations
- They should be comparable (up to some point)

Archetype	Elements†	Display	Avatar View	World View	Action*
Face-to-face Window	F	2D/3D Screen	2D/3D Video	N/A	-
Shared VR	M	VR HMD	CGI	CGI	O, P, L (virt)
Remote Asssistance	V	VR HMD	N/A	2D/360 Video	P (twin/phy)
Social XR	F, M	HMD	2D/3D Video	CGI	O, P, L (virt)
Immersive Digital Twin	M, V	VR HMD	CGI	3D Photo	O, P (twin)
Immersive Telepresence	V, F	HMD	360 video		O, P, L (phy)
Distributed Reality	F, M, V	Concept Only			
AR Host	M (simple)	AR HMD	CGI cues	N/A	P (twin)
Audio conference	-	Screen	Icons	-	-

†F(ace), V(isit), M(eet).
*O(bject), P(ointer), L(ocomotion); phy(sical), virt(ual), twin.

Overview of existing works

Work	Archetype	Context	Dialogue	Exploration	Manipulation	Conditions	High Level feature	Low Level Feature
Kim et al. (2019b)	Face2Face Window	Watch a movie	Comment on the movie	–	–	Big display / small display with following gaze / corner display	Subscale from NMM-SPI for emotions Likes/dislikes about gaze following	–
Lawrence et al. (2021)	Face2Face Window	Conferencing	Semi structured conversation with a research confederate	–	–	3D video conferencing vs. 2D video conferencing system	HOLO for presence, attentiveness, connectedness, reaction-gauging	HG, HN, and EM for non-verbal behaviour
Pan and Steed (2017)	Shared VR	Play games	–	–	Solving puzzles (pieces)	Embodiment types: no self-avatar, self-avatar, and face2face	IT for trust	CT for task performance
Li et al. (2019)	Shared VR	Photo sharing	Comment on shared photos	–	–	Face2face, skype, and Facebook Spaces	SocialVR for social presence PMRI for emotions	–
Orts-Escolano et al. (2016)	Social XR	Play and collaborate remotely	Tell a lie game and dialogue to build blocks	–	Building blocks	VR vs. AR	Semi-structured interview for: presence, interaction, explorability, etc.	Semi-structured interview for: visual quality and latency
Gunkel et al. (2018)	Social XR	Watch a movie and play a game	Comment on the movie or game	–	–	System performance	RGQ: Social presence, interaction, explorability, and global QoE	RGQ: visual and audio quality
Prins et al. (2018)	Social XR	Play a game	–	–	Pong (two-player game)	System performance	Feedbak on presence and overall quality	–
Lawrence et al. (2018)	Social XR	Conferencing	Negotiation	–	Assembly of lego blocks	Audio only, video fixed to HMD, and video fixed to host world	NMM-SPI	–
Li et al. (2021)	Social XR	Watch a VR Movie (inside the scene)	Questions raised by movie characters	Follow Characters	Interact with environment (e.g., click buttons)	HMD vs. screen with game controller	WS for presence SocialVR for social presence SSQ for cybersickness NASA-TLX for mental workload	VQoE for visual quality
Lee et al. (2018)	Remote Assistance	Remote collaboration	–	Find a set of target objects in the task space	Place the target objects on the desk	Dependent view vs. independent view	MEC spatial for presence NMM-SPI for social presence SMEQ for workload SSQ for cybersickness	CT for task performance
Young et al. (2019)	Remote Assistance	Remote exploration	–	Explore remote environment	–	Three ways of interaction	IPQ for presence NMM-SPI for social presence and emotions SSQ for cybersickness	–
Teo et al. (2019a)	Remote Assistance	Remote collaboration	–	Identify objects	Decorate a bookshelf placing objects	No cues, hand gestures, pointer, and hand gestures + pointer	MEC for spatial presence NMM-SPI for social presence SMEQ for workload	SUS for usabilty
Teo et al. (2020)	Remote Assistance	Remote collaboration	–	Find a set of target objects in the task space	Place the target objects to a specific location	360 image mode vs. 360 projection mode	MEC for spatial presence NMM-SPI for social presence SSQ for cybersickness SMEQ for workload SEQ + 3 custom questions for global QoE	SUS for usabilty CT for task performance
Wang et al. (2020)	Remote Assistance	Remote collaboration on physical tasks	–	Locate blocks and follow remote pointers	Assembly of lego blocks	Cursor pointer, head pointer, and eye-gaze pointer	TQ: co-presence, interactivity, and explorability NASA-TLX for mental workload	TQ: Visual and audio quality
Bai et al. (2020)	Remote Assistance	Work together remotely	–	Search blocks and follow remote indications and visual cues	Assembly of lego blocks	Verbal only, eye gaze, hand gesture hand gaze + hand gesture	MEC for spatial presence NMM-SPI for social presence NASA-TLX for mental workload	SUS for usabilty CT for task performance
Anjos et al. (2019)	AR Host	Play games.	Solve riddles	–	–	System performance	–	Semi-structured interview for task performance
Kasahara et al. (2017)	Immersive Telepresence	Cleaning up a lab room	–	Locate objects and clean them	–	Video Stabilization	SSQ for cybersickness	HM for non-verbal behaviours
Piumsomboon et al. (2019)	Immersive Telepresence	Remote collaboration	Guess objects of interest	House inspection	Arrange objects	Types of virtual representations, levels of miniature control, levels of 360-video view dependencies, and 360-camera placement positions	MEC for spatial presence NMM-SPI for social presence SSQ for cybersickness SMEQ for workload	SEQ for task performance
Zhang et al. (2020)	Immersive Telepresence	Telepresence	–	Locate / indicate remote user's gaze	–	Distance to avatar and display (AR, tablet)	–	AE for task performance
Piumsomboon et al. (2018)	Immersive Digital Twin	Remote collaboration	–	Identify objects	Place objects to a specific location	Fixed life-size full-body avatar with and without Mini-me	NMM-SPI for social presence SMEQ for workload	SEQ for task performance
Brunnström et al. (2020)	No communication	Remote control	–	–	Control a crane to load logs	Latency	RC: Comfort, immersive, and overall quality	RC: Picture, responsiveness, and task accomplishment quality
Pérez et al. (2021)	No communication	Escape room game	–	–	Manipulate game objects	Real hands vs. VR controllers	WS for presence Embodiment DREQ for Global QoE	–

In summary

P.1320

- Test design
  - System IF: independent variable
  - Human IF / Context IF: controlled
  - QoE constitutents: dependent var
- Two objectives
  - Systematically control independent variable
  - Test complete black box systems
- Tasks?

SoA Analysis

- System classification according to perceptual characteristics
- 3 families of tasks
  - Deliberation
  - Exploration
  - Manipulation

Proposed target

- Identify 3 simple tasks: deliberation, exploration, manipulation
  - Feasible in all system archetypes
  - Potential to react to SIFs and influence in QoE constitutents
  - **Do not cover all use cases: maximize coverage with only 3 tasks**
- Define a **basic (but specific) test methodology** around these tasks
  - How to consider/control HIF/CIFs
  - How to structure test session
  - ...
- Validate / explore the test methodology with a cross-lab test
- Liaison with ITU-T SG12

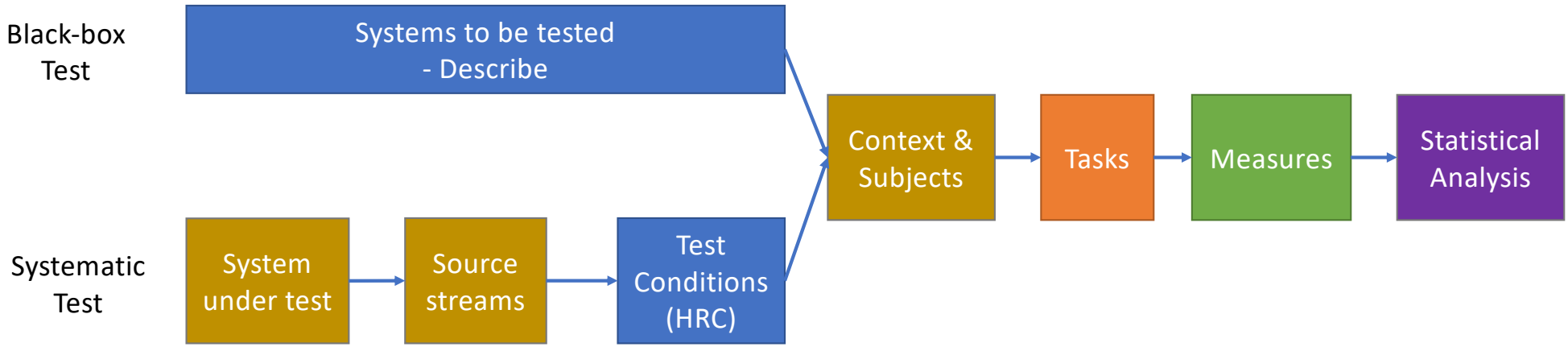
Proposed approach to cross-lab test

- All labs agree on
  - Tasks to be executed (same in all labs)
    - But we could try different tasks to select among them
  - Basic test methodology
  - Which QoE constitutents to evaluate and how (??)
    - To be decided: everyone uses same measures? Should we provide a subset / superset?
  - How to report test results
- Each lab proposes / decides
  - Which system(s) to test
  - Under which conditions (SIFs)
  - The actual context of the experiment (e.g. use case, setting…) in a way which makes sense for them

Test methodology

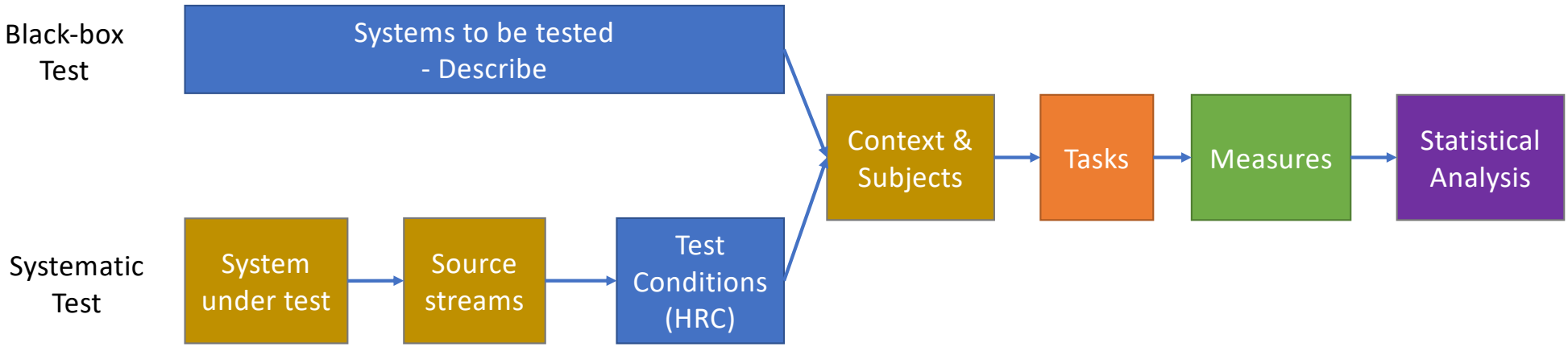
Hypothesis

Variations in the **independent variable** have **significant effect** on a **QoE Constitutent** when user execute a basic **communication task** under controlled **conditions**



Test methodology

- Define a common way to describe test **conditions** and **variables.**
- Standardize basic **communication task.**
- Standardize a few measures (e. g. behavior analysis) that can be applied to many tests and systems.
- Offer appropriate **statistical analysis methods** to describe measure results and its effect



Systems to be tested

Affiliation	System Name	Developed by	Description (receiver view)	FACE: how is the remote person represented?	MEET: how is the shared space represented?	VISIT: how is the environment of the remote person represented?	MOVE: How is your body represented? Which type of interaction do you have?
Nokia XR Lab	The Owl (remote)	Nokia XR Lab	I wear an HMD (Meta Quest) and see a remote place in 360 video. Other remote users are seen as avatars	Cartoonish avatar	360º video + avatars	360º video	Own hands using video passthrough and segmentation
Nokia XR Lab	The Owl (local)	Nokia XR Lab	We are at a meeting room with a 360-degree camera and remote people attend the meeting	Cartoonish avatar	N/A	N/A	N/A
CWI	VR2Gather	CWI	I can access through desktop or through HMD. I am surrounded by 3D cameras capturing my likelihood in 3D. I see a virtual world, with other user represented as volumetric contents.	Real-time volumetric capture	Synthetic volumetric objects	N/A	Body represented through volumetric video. Can interact through controllers.
UPM	FVV Live (remote)	UPM	I can access through smartphone, desktop or HMD and see a volumetric representation of a remote place	Real-time volumetric capture	Volumetric representation	N/A	TBD
UPM	FVV Live (local)	UPM	We are sourrounded by cameras capturing us and the environement	TBD	TBD	N/A	Volumetric video

- **More** (including commercial systems)**?**

Communication tasks

- **Deliberation:** Conversations between peers, normally oriented to achieve a common goal.
- **Exploration:** Exploration of the environment and identification of objects following indications.
- **Manipulation:** Interaction with system elements and manipulation of physical objects (e.g., lego blocks).

Communication tasks

Deliberation	Exploration	Manipulation
Commenting on shared content (videos, photos, games,…)	Identifying and locating objects in the task space	Manipulating/placing physical objects (e.g., Lego blocks)
Follow instructions (e.g., physical exersices)	Exploration of the remote environment	Interact with virtual objects (e.g., press buttons, objects used in games (e.g., puzzles), …)
Play games (tell-a-lie, riddles, etc.)	Follow characters	
Semi-structured interviews		
Negotiation		

Tentative plan

- Step 1 (dec'22 – may'23): planning
  - Start **regular audio calls**
  - Call for systems / participants
  - Decide on common elements (methodology, result report)
  - VQEG F2F: Each lab presents their experiment proposal. Close list of participants.
- Step 2 (may'23 – dec'23): execution
  - Each lab executes the test
  - (Common) Results are shared in agreed format
  - VQEG F2F: Present and discuss individual results.
- Step 3 (dec'23 – may'24): evaluation
  - Evaluation of the test methodology based on results → common paper, ITU contribution
  - Evaluation of each individual experiment → each lab will exploit their results

Potential participants

- Identify interested people/groups:
  - UPM (Spain)
  - Nokia (Spain)
  - CWI (The Netherlands)
  - RISE (Sweden)
  - TU Ilmenau (Germany)
  - Ghent University (Belgium)
  - Wuhan University (China)
  - University of Surrey (UK)
  - AGH University of Science and Technology (Poland)
  - University of Brasilia (Brazil)
  - University of Padova (Italy)
  - University of Roma 3 (Italy)
  - Nantes Université (France)
  - UCL (UK)
  - Athlone Institute of Technology (Ireland)
  - **You?**