VQEG

IMG Test plan on Immersive communication systems

2023-12 – VQEG Plenary Meeting (Online)

Pablo Pérez (Nokia, Spain)
Jesús Gutiérrez (Universidad Politécnica de Madrid, Spain)
Kamil Koniuch (AGH University, Poland)
Ashutosh Singla (CWI, Netherlands)

Motivation

- How to test a bi-directional immersive communication system:
  - Evaluate effect of technical factors in QoE (e.g. variations of latency / bitrate / etc.)
  - Compare with other systems / experiments
- Background:
  - **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.1301** - Subjective quality evaluation of audio and audiovisual multiparty telemeetings
  - **ITU-T P.1320 (P.QXM)** - QoE Assessment of eXtended Reality (XR) Meetings
    - Best practices, factors and constituents for QoE assessment of telemeetings with XR elements.
    - Not to the detail of proposing evaluation tasks or methodologies.
  - Overview, Taxonomy and Good Practises (article*)
    - Characterize types of systems and 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

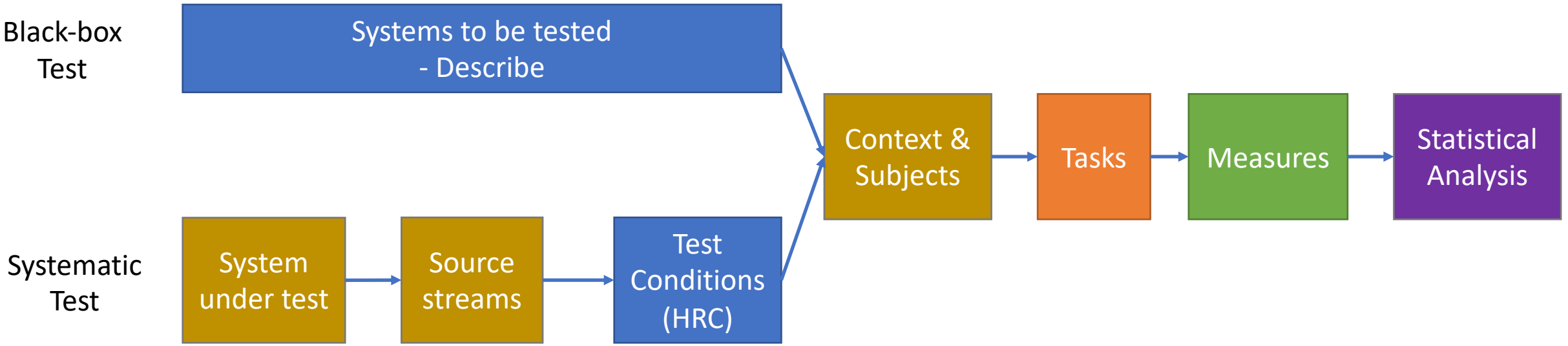
Work item ITU-T P.IXC

- VQEG-IMG and ITU-T SG12 are working on the definition of a new **recommendation** for subjective assessment of eXtended Reality (XR) communications:
  - A **methodology** to describe the test design: which system influencing factor to test, how to control context and human influencing factors, which QoE constituents to address. The methodology should cover two types of designs:
    - systematically control an independent variable and observe the effects on QoE.
    - test complete “blackbox” systems without exploring individual variables.
  - A reduced set of communication-based **interactive tasks** that are suitable for testing XR in communication systems.
    - 4 tasks: audio communication, visual communication, object manipulation, and environment exploration.
    - Not covering all possible use cases or XR systems but maximizing the coverage provided by only a few tasks.
  - A subset of relevant **measures**: behaviour analysis, questionnaires regarding QoE constituents, physiological measures
    - Objective is to recommend a few measures which can be applicable to a wide range of use cases and systems

Methodology

Hypothesis

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



Communication Systems

- Immersive communication systems:
  - Enable immersive, remote, and synchronous communication.
  - Immersive systems are characterized for creating the illusion of “being there” (a “sense of presence”), in a remote location and/or with remote people.
- Target systems that has at least one of the following features:
  - Transmit in real time a realistic visual representation of the other person (**Face***): This element enables visual communication.
  - Transmit in real time a visual representation of the surroundings of the other person (**Visit***): This enables remote presence: seeing the physical environment of the other person and being able to operate and discuss it.
  - Represent the other person in the same (virtual or physical) space as the user (**Meet***): This enables shared immersion.

* 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

Tasks

- **Audio**-communication task: Survival Task
- **Visual**-communication task: Charade / Physiotherapy
- **Object**-based communication task: Block building
- **Environment**-based communication task: Treasure Hunt

Call for participation

- **Objective**: To look for laboratories which would like to perform subjective assessment tests based on the protocol defined for the Recommendation, so that we can validate the method itself before proposing it to ITU-T.
- Published in May 12th, 2023.
- Important dates and (tentative) schedule:
  - May 23, 2023: Declaration of interest.
  - June 16, 2023: Submission of the form and a 1 to 3-slide presentation describing your test.
  - June 29, 2023: Presentation of the slides in VQEG F2F Meeting (Hybrid).
  - July – December 2023: Each lab executes the test and (common) results are shared in agreed format
  - December 2023 (TBD). Presentation of the results in VQEG F2F Meeting (Virtual).
  - December 2023 – May 2024:
    - Evaluation of the test methodology based on results → common paper, ITU contribution
    - Evaluation of each individual experiment → each lab will exploit their results.

Overview
Labs, tasks, and systems

Task	#Labs	Systems				
		Face2Face Window	**SocialVR**	**SocialXR**	**Immersive Telepresence**	**2D Baseline**
Survival (audio)	**AGH**, TUI UPM,NOK, TSU, KUM	FVV	TUI-SocialVR, Mozilla Hubs	VR2Gather	Owl	Teams
Charade / Physiotherapy (visual)	**CWI**, KUM, TSU, USU, LUT, UPM	FVV	Mozilla Hubs	VR2Gather		Teams
Block building (object)	**UPM**, UR3, UCL, LUT	FVV	UR3/PAD-Social VR, UBIQ	VR2Gather		Teams
Treasure Hunt (environment)	**NOK**, PAD, PAD, RISE, ERI+UCB	RISE-RemoteControl	UR3/PAD-Social VR		Owl DualStream	

13 Labs!

Overview
Progress from July 2023

- Alignment on:
  - Task definitions → (Almost) Completed!
  - Measures (e.g., questionnaires) → Under discussion.
  - Tests without any technology (baseline) → Agreed (in some tasks?).

Task: Survival
Audio-communication

- Based on the one described in the **ITU-T Rec. P.1301**
- **Aim**: To have comparable times of engaging discussion.

- **Difficulties with original game**:
  - Time of discussion heavily dependent on human factors.
  - Feedback is unsatisfying for participants

Task: Survival
Audio-communication

Changes provided after pre-tests:

- Game iterations: 5, 10, 15 items.
- Pretest: 5 items without technology sets the main test scale.
- Objective: List possible uses and arguments for each item.
- Feedback: Examples provided for item applications.
- Discussion time: 6 minutes with a time-alerting clock.
- VR integration: Aim to keep measurements and instructions in VR.
- VR representation: Items shown as boxes with 2D pictures and descriptions.

Task: Survival
Audio-communication

Lab	Systems	Participants	SIFs	Type	Research Question
AGH	Live / Teams / VR2Gather	3-5	[Systems]	Blackbox	Systematic evaluation of the task
UPM	Live / FVV	3 (1+2)	Display	Systematic	Systematic evaluation of the system parameters
TUI	Live / SocialVR	3	Live vs. audio spatialization	Blackbox	Block/blackbox comparison of 2-3 systems/conditions

The Effect of Audio Spatialization on social presence and user behavior in multi-party VR Communication

VQEG-IMG

Felix Immohr, Alexander Raake

Audiovisual Technology Group, Institute for Media Technology
Technische Universität Ilmenau

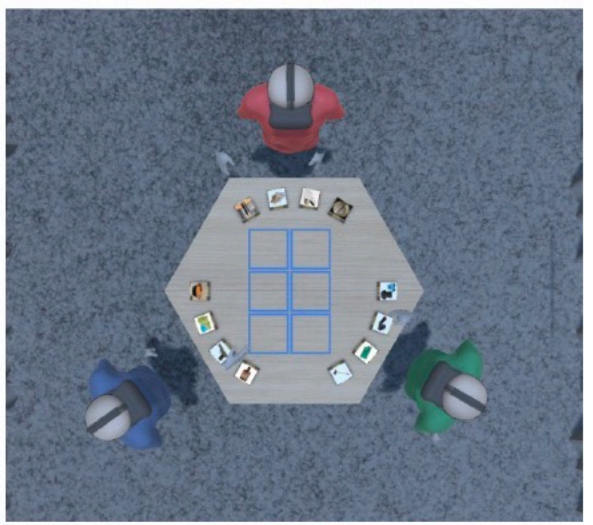
Audiovisual Technology Group

The SPIRIT of science

TECHNISCHE UNIVERSITÄT ILMENAU

Study Design

- Interactive audio-visual communication test with 3 participants
- **Research Question:**
  Effect of spatial audio reproduction on social presence and user behavior in multi-party VR communication?
- **System factors** (independent variables):
  - Auralization method (diotic, position-dynamic binaural synthesis)
  - ‚Communication medium' (VR, face-to-face)
- **3 Conditions, within-subject design:**
  - VR with **binaural spatial audio**
  - VR with **diotic audio**
  - Face-to-face interaction



Audiovisual Technology Group

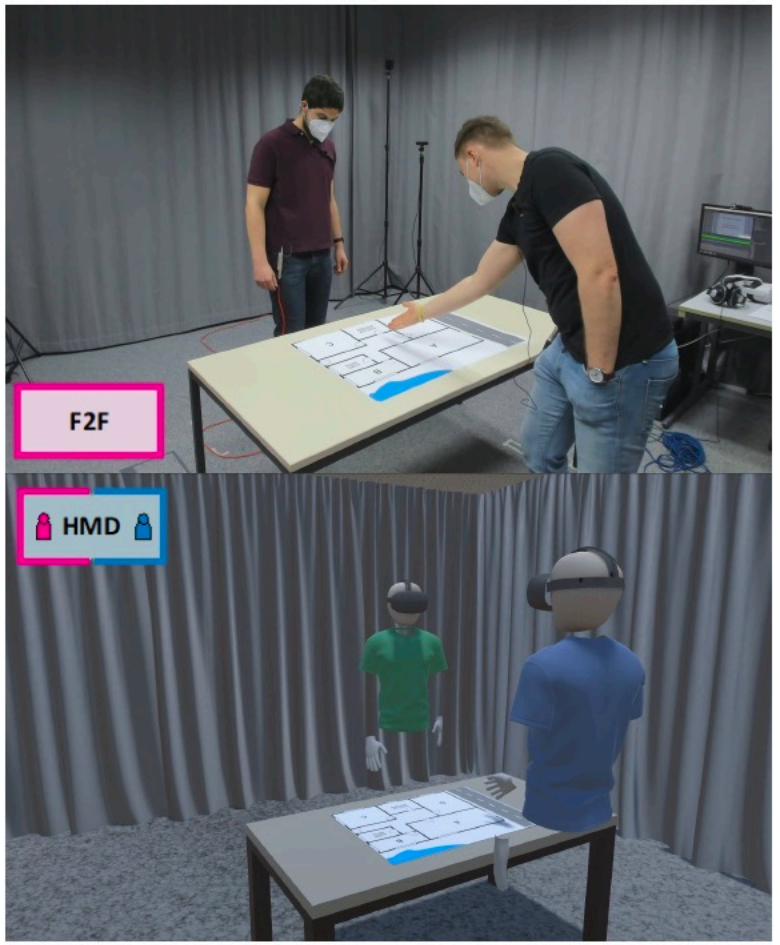
The SPIRIT of science

TECHNISCHE UNIVERSITÄT ILMENAU

Study Design

- **Evaluation** (dependent variables):
  - **Direct** assessment through questionnaires:
    - Social presence
    - Audiovisual plausibility
  - **Indirect** assessment: conversation and behavioral analysis
    - Conversation analysis (turn-taking behavior)
    - Behavioral analysis (movement)

- Multi-party extension of [1] with adapted survival task

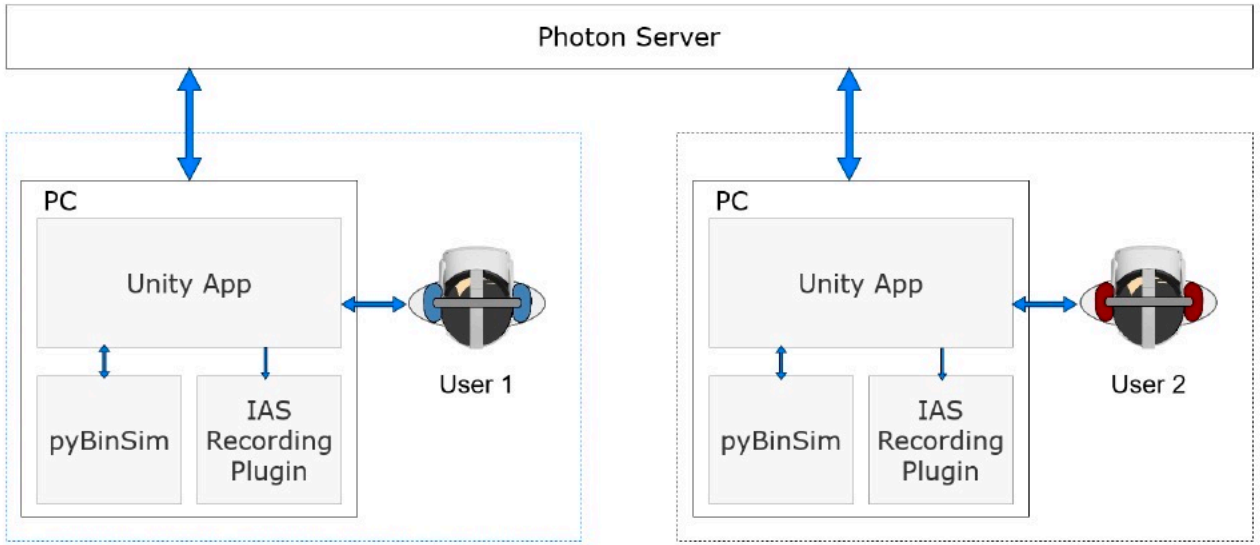


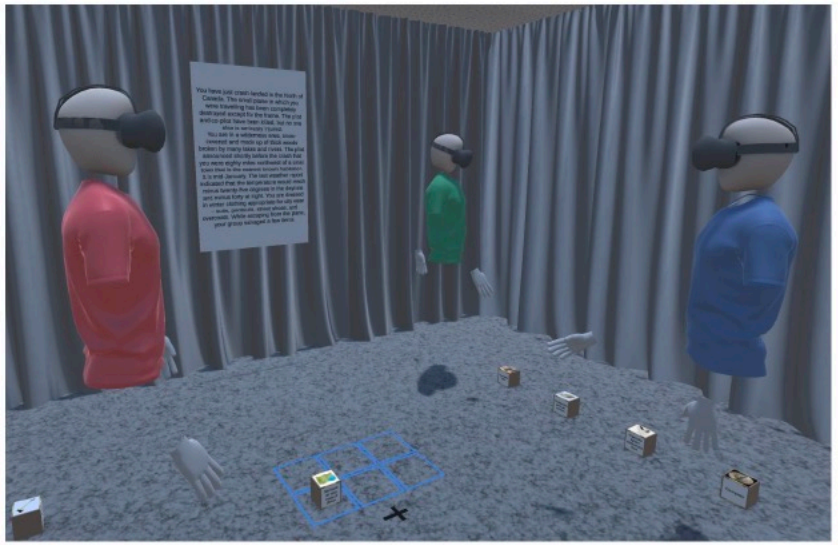
Example study scene

[1] Immohr, F., Rendle, G., Neidhardt, A., Göring, S., Rao, R. R. R., Arboleda, S. A., Froehlich, B., Raake, A., „Proof-of-Concept Study to Evaluate the Impact of Spatial Audio on Social Presence and User Behavior in Multi-Modal VR Communication", ACM Int. Conference on Interactive Media Experiences, Nantes, France. | https://doi.org/10.1145/3573381.3596458.

Study Execution and Analysis Framework

- Audiovisual virtual environment for conducting multi-party communication studies
  - Audio and scene state distributed to each user
  - Binaural audio reproduction: pyBinSim (Neidhardt et al., 2017)
  - Experiment control: bmlTUX (Bebko and Troje, 2020)
  - Session recording and playback with Unity-based plugin





Neidhardt, A., Klein, F., Knoop, N. and Köllmer, T., "Flexible Python tool for dynamic binaural synthesis applications", 142nd AES Convention, Berlin, 2017.
Bebko, A. O., Troje, N. F., "bmlTUX: Design and Control of Experiments in Virtual Reality and Beyond", In: i-Perception 11(4): 1-12, 2017. | https://doi.org/10.1145/3573381.3596458

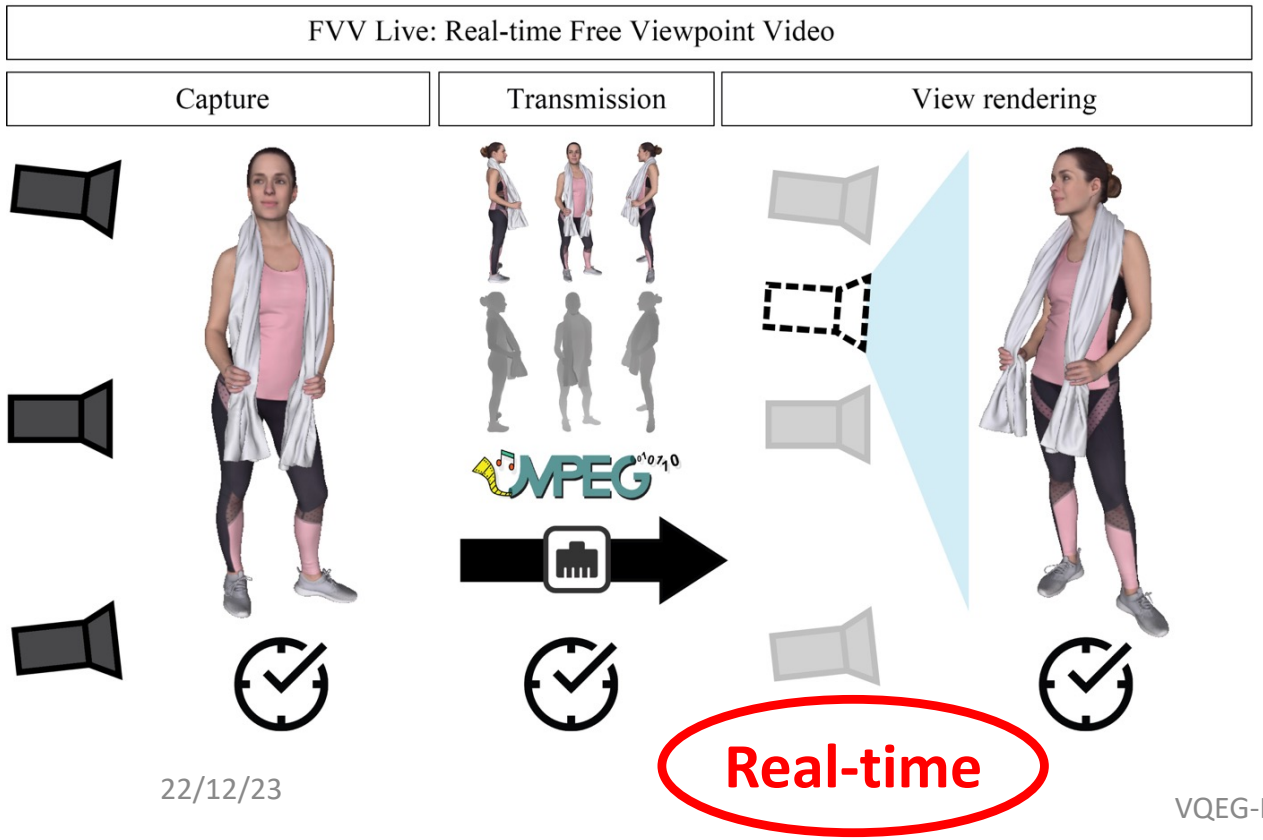
Universidad Politécnica de Madrid

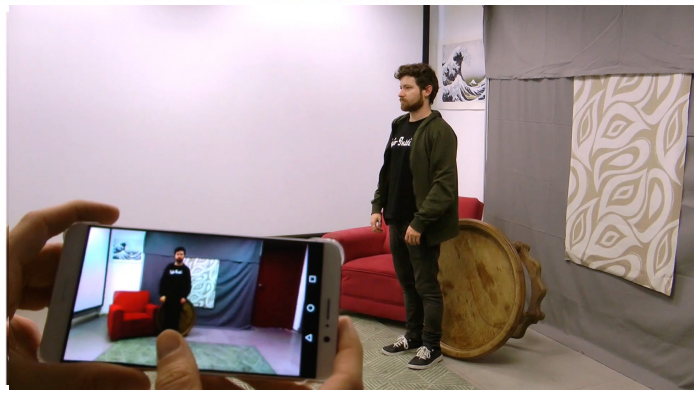
Jesús Gutiérrez – jesus.gutierrez@upm.es
Carlos Cortés – carlos.cs@upm.es
Miguel Die – m.die@alumnos.upm.es
Narciso García – narciso.garcia@upm.es

Task: Survival
Audio-communication

UPM Proposal

- **RQ:** How technical parameters of the **FVV Live system** impact the QoE of the sueres during communication scenarios?





Additions:
- Bidirectional audio
- Representation of the remote user through a video on a tablet.

Task: Survival
Audio-communication

VQEG

UPM Proposal

SIF
(Independent variables)

Display technology (HMD vs. screen+joystick)

QoE factors
(Independent variables)

Satisfaction with the discussion → MOS (5 levels)
Overall QoE → MOS (5 levels)
Audiovisual quality → MOS (5 levels)
Communication easiness → MOS (5 levels)
Social and spatial presence → Marta [1], [2]
Interaction → QoI [3]
Subjective performance → MOS (5 levels)
Task load → NASA TLX
Cybersickness → SSQ and Vertigo scale [4]
Behavior → Audio recordings
Task performance → Comparison with solutions

Participants

- 3 users each time (one remote and 2 local)
  - Total number of participants: 54

Test session

1. Training and pre-questionnaires
2. Test without technology
3. 3 Tests with one in remote with HMD/screen.
   - 5 items

Each test: Agree on the main 2 applications of each item

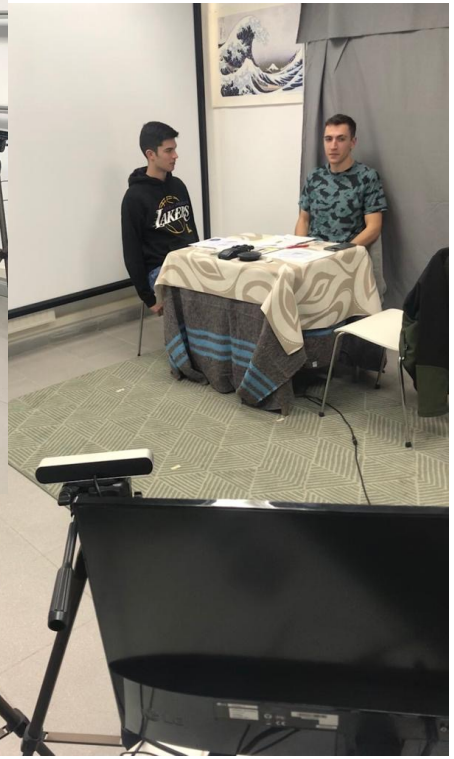
- 2 minutes for individual thinking of the applications (clock)
- 6 minutes for discussion (clock)
- 5 minutes for post-questionnaire

[1] M. Orduna M et al.. "Methodology to assess quality, presence, empathy, attitude, and attention in 360-degree videos for immersive communication", IEEE Transactions on Affective Computing. 2022 Feb 14.
[2] B.G. Witmer and M.J. Singer. "Measuring presence in virtual environments: A presence questionnaire". In: Presence 7.3 (1998).
[3] K. Gupta, G. A. Lee, and M. Billinghurst, "Do you see what i see? the effect of gaze tracking on task space remote collaboration," IEEE Transactions on Visualization and Computer Graphics, 2016
[4] P. Perez, N. Oyaga, J. J. Ruiz, and A. Villegas, "Towards Systematic Analysis of Cybersickness in High Motion Omnidirectional Video", QoMEX 2028.

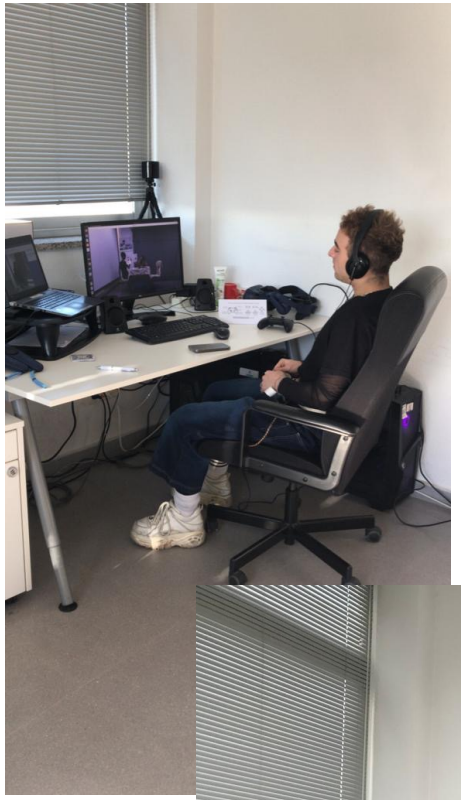
Task: Survival
Audio-communication

UPM Proposal

- Tests going on → Results available soon!



Local users



Remote user

Task: Charade / Physiotherapy
Visual-communication

- **Goal**: to perform a task that involves visual-only, or visually-predominant, communication from the participants.
- **Examples:** using sign language to perform the conversation, or conducting a physical training session.
- Depending on the DoF (i.e., how free they are to move around the space), we define **two subtasks**:
  - 3DoF (the participants are seated and can move their upper body): Charade.
  - 6DoF (the participants are free to move around the space): Physiotherapy training session
- The success is measured by the number of times the task is performed (words or movements) in the fixed amount of time.

Task: Charade / Physiotherapy
Visual-communication

- **Physiotherapy training session**: Participant A is the instructor, who must demonstrate a finite set of moves that Participant B must replicate correctly. The session ends when Participant A inspects Participant B and agrees that the moves have been performed correctly. The success is measured in terms of time taken to complete one move, or on moves completed per time unit.
- **Charade**: Participant A is given a list of words that participant B must guess, in a finite amount of time; to aid the guessing, participant A cannot use oral instructions, but must rely on physical gestures. The game ends when all words have been guessed, or when the time is up. The success is measured in terms of words completed per time unit

Task: Charade / Physiotherapy
Visual-communication

Lab	System	Participants	Experiment	Research Questions	Timeline
Texas State University	2 VR System + MS Teams	2	Charade	Performance of 2 VR systems will be compared with MS Teams	Feb/March-end
University of Surrey	MS Teams + VR System	2	Charade	Investigating the imapct of Packet Loss on a Teleconferencing Software	April-end
Keysight	MS Teams	2	Physiotherapy & Charade (TBD)	Impact of radio and network impairments on telemeeting QoE	Spring 2024
CWI	VR2Gather	2	Physiotherapy	Impact of Latency and Desynchronization on XR Communication	March-end

EXPLORING AUDIO-VIDEO SYNCHRONIZATION THRESHOLDS FOR ENHANCED IMMERSIVE EXPERIENCES IN XR COMMUNICATIONS

A. Singla, I. Viola, J. Jansen, P. Cesar
CWI, Netherlands

Research Goals and Task Description

Research Question

- Impact of Latency and Desynchronization on XR Communication

Task Description

- Confederate user acting as a physiotherapist in the experiment
  - Teaching role involving demonstration of various exercises

- Participants required to replicate exercises shown by the confederate user
  - Provision of feedback on correct or incorrect exercise performance
  - Guidance offered by confederate user for correcting exercise techniques
  - Progression to next exercise upon satisfactory performance

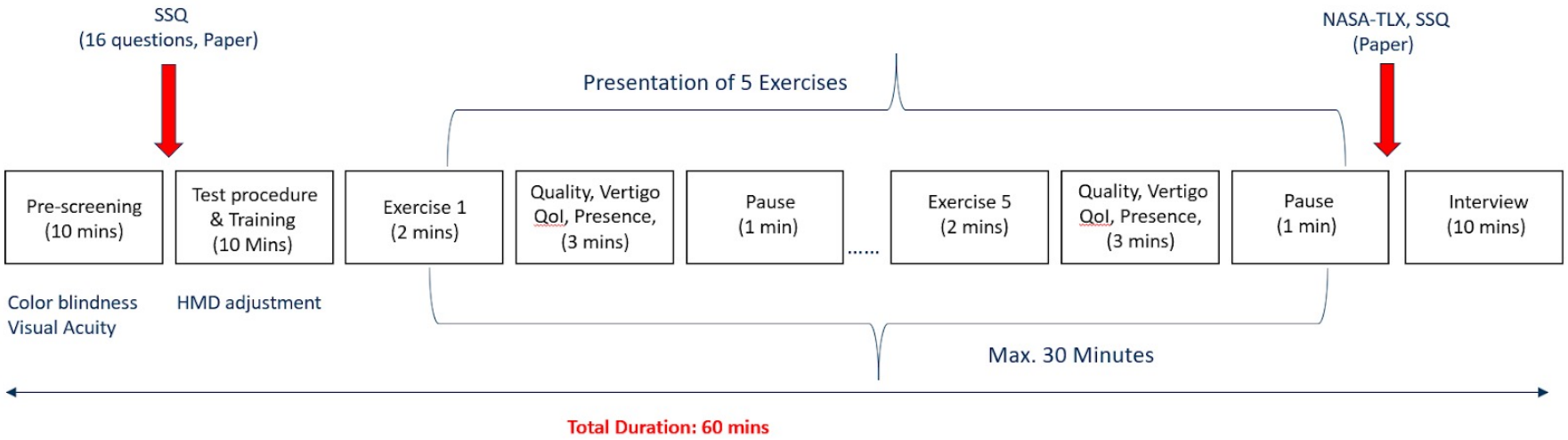
- Sequential learning of exercises throughout the experiment

Technical Setup

- **System Factors:** Latency will be introduced in Audio and Video
  - Finding acceptable limits of delay value for audio and video for smoother and immersive XR experience
- **System**
  - VR2Gather [1] will be used for carrying out this experiment and will be modified based on the requirements
  - System latency will also be measured before the start of the experiment and during the experiment as well
- **Encoding**
  - No encoding will be applied. The confederate user and participants will see each other in the highest available quality
- **Test Design:** Within-subject
- **HMD:** Quest Pro

[1] I. Viola, J. Jansen, S. Subramanyam, I. Reimat and P. Cesar, "VR2Gather: A collaborative social VR system for adaptive multi-party real-time communication," in IEEE MultiMedia

Test Method



- **Data Collection**
  - Visual Quality
  - Quality of Interaction
  - Presence/Immersion
  - Cybersickness
  - Task/Cognitive Load
  - Task Completion Time
  - Head Rotation Data (Yaw, Pitch and Roll)
  - Movement Data (X, Y, and Z)

https://docs.google.com/document/d/1MSx663dEyB1JqPRtkvAyY1YGa_sMZHHqpLBf_ie4jAg/edit

Labs & Test Set Up

MediaScape: https://youtu.be/I7kY1cMZyD0

KEYSIGHT

VQEG IMG project

M. Lecci, A. Pérez Aguilar, A. Díaz Zayas, G.C. Madueño

Research Question

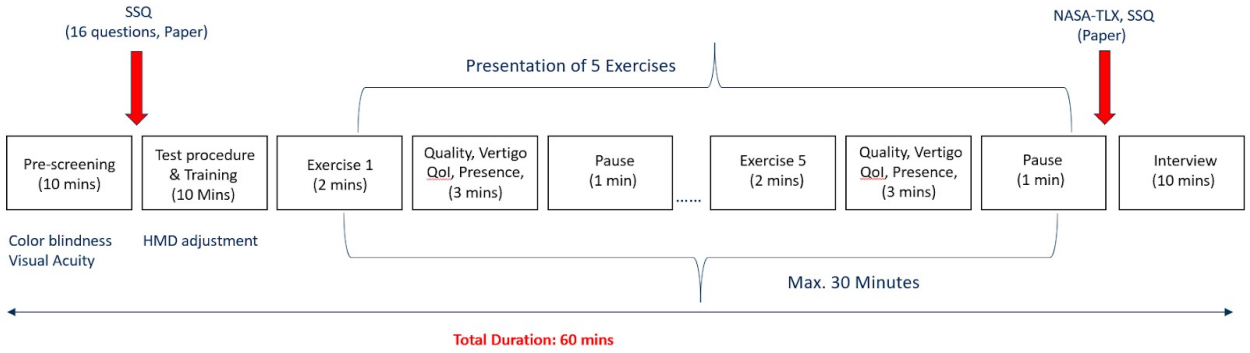
- How do radio and network impairments affect the telemeeting QoE?
  - Setup 1 using Y.TestBed scenarios: How do telemeetings are affected by typical real-life scenarios?
  - Setup 2 using tc-netem: How do telemeeting are affected by ideal network impairments?

Experimental system

- First time dealing with interactivity in our testbed
  - Common traits: Microsoft Teams using laptops (TV or projector might be an option for better immersiveness)
  - Setup 1: connection coming from controlled testbed (Y.TestBed), with all radio and LAN parameters under full control
  - Setup 2: over-the-air connection from indoor private 5G pico-cell, with partial control over radio parameters (wireless connection is inherently uncontrollable) and full control on LAN parameters
- Setup 1 is our target, although we still lack a solution to make it work reliably
- Setup 2 is our fallback simpler option

Test design

- **Devices**: laptops, possibly connected to large TVs or projectors
- **Meeting software**: Microsoft Teams
- **Experiment**: Physiotherapy (6-DOF). If time allows and students participate, also Charade (3-DOF)
- **Representation**: webcam video + laptop microphone
- **Environment**: meeting room
- **Data collection**: screen recording, audio recording, PCAP traces, core network and radio parameters (depending on the system used), Questionnaire (same as other groups)



Time plan

- Pilot testing
  - February-March 2024
  - 2-4 subjects from the group
  - Test feasibility for setup 1&2, decide which one to use
- Actual test and data analysis
  - Spring 2024
  - ~1 month for human panels (depending on student participation)

Realistic Representations or Virtual Reality: Which system is better for communications systems?

Mylene Farias, mylene@txstate.edu

December 2023

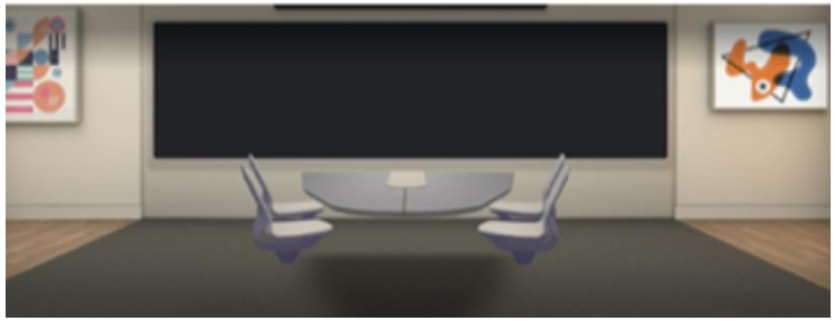
Research Goals

- **Research Questions**
  - Compare the performance of 2 VR systems and a baseline Teleconference systems
  - Which system provides a better:
    - Engagement, presence, comfort, performance
- **Task Description**
  - 2 participants playing a charade game
  - 3 different setups: 2 VR systems (Quest2) and MS Teams
  - No voice, only gestures, seated participants

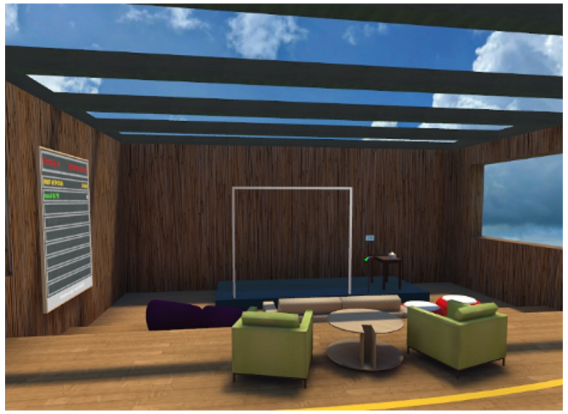
Task	Feature	Participants	DoF	Performance / success measure	Beyond VR	Confederate	Symmetric
Charade	Visual	2-10+	3DoF	words/time	-	No	Yes

Illustration of Systems

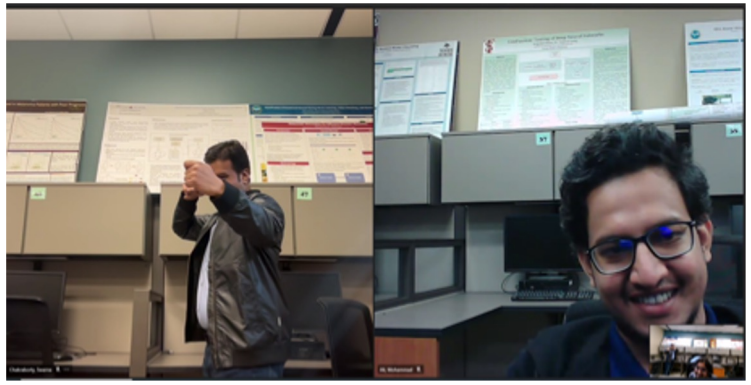
Horizons

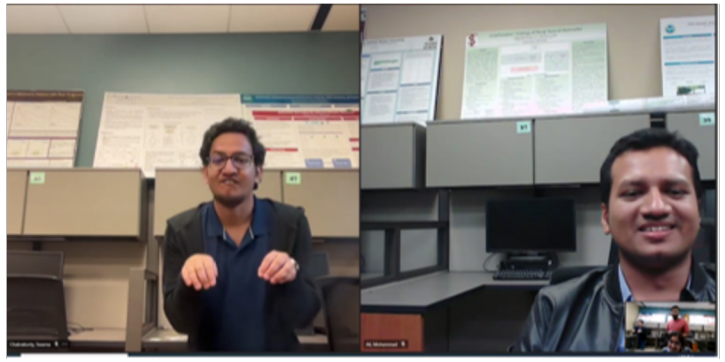


Chat VR



MSTeams





Test Method

- Performed a **Pilot Test (December'23)**
  - 5 x 2 participants in each system
  - A list of simple predefined words
    - Experimented with simple words separated in classes
    - All participants have English as a second language - something we considered
  - Participants alternated as guesser and mime, within a fixed amount of time (tried 10 and 2 minutes)
  - Preliminary Statistical analysis is not ready yet
- Pre and Post questionnaires only
  - Should we consider an intermediate question?

Test Method: Data Collection

- Pre-questionnaire
  - Personal information
  - Experience with charade game, VR/AR and HMDs
  - Motion/Cyber sickness questionnaire (R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, “Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness,”)
- Post-questionnaire
  - Motion/Cyber sickness questionnaire
  - Presence questionnaire
  - Task Load
- Some comments:
  - Lack of diversity among users and non-native speakers
  - Hard words were difficult
  - Hand tracking only works for hands in the view of the hmd. Larger movements are not possible.

Technical Setup

- VR Systems : Quest2
  - Hand tracking ability was used so that avatars could mimic the words
  - In Horizons, a small window in the display allowed participants to see written words
  - In VRChat, words were passed through chat
- Monitor Displays were used for MS Teams
  - Words were passed/shown in papers
- We were not able to capture eye-tracking or head tracking data at this time

Next Steps

- Get questionnaires approved
  - Are there too many questions?
  - What other data should we record?
- Perform experiments in mid January to February when classes restart

MEASURING THE IMPACT OF PACKET LOSS BLACK-BOX MIXED REALITY COMMUNICATION

A. ADEYEMI-EJEYE

UNIVERSITY OF SURREY, UK

RESEARCH GOALS

- **Research Question**
  - What level of packet loss is required to impair mixed reality communication usinf a black-box teleconferencing software?

- **Task Description**
  - Subjects will play a 2- person charade game in a Mixed reality environment with the communication via MS Teams. The methodology will include a pre-questionnaire, a training, and a post questionnaire. Point 2
    - One party will have a camera and screen, while the second party will use a HMD to receive the visual information. Both parties will also be able to communicate via audio

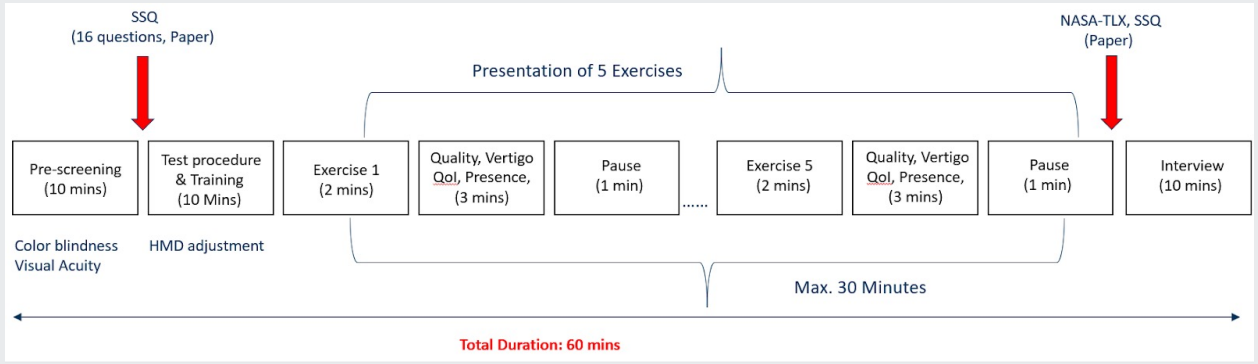
- The charade exercises have to be in a fixed order. The exercises should also be arranged in a sequence of increasing difficulty

TECHNICAL SETUP

- **Test System**
  - MS Teams
  - Immersed
- **Encoding**
  - Native MS Teams encoding and rate adaption
- **Test Design**
  - Ms Teams using laptops
- **HMD or 2D screen**
  - Both (one party uses a 2d screen and the other with use a HMD

- Any other point you may consider to highlight
  - We would also be dealing with packet loss mitigation solutions within teams so we might have to apply higher levels of packet loss to see the impact

TECHNICAL SETUP



- **Data Collection**
  - Quality
  - Simulator Sickness
  - QoI
  - Head Rotation Data (Yaw, Pitch and Roll)

TIME PLAN

- **Pilot Testing**
  - Mid – End of January
  - 4 participants
    - Feedback would help amend the final test

- **Actual Test and Data Analysis**
  - Beginning of February – End of March (Test)
  - April (Data Analysis)

- Open Questions
  - Where might be best to apply Packet Loss (Which participant)

Task: Block building
Object-based communication

- **Goal**: to perform a task that involves audio-visual communication for remote collaboration that implies the manipulation of objects.
- Similar tests already performed with the VR2Gather system (collaboration CWI-UPM).
- Users will be able to see representations of themselves, other users and the blocks to build with.
- Users will be able to manipulate the blocks in two ways:
  - **Instructor**: will be able to manipulate a construction composed by blocks.
  - **Builders** (rest of the users): have to replicate the construction of the instructor following his/her verbal indications.

Task: Block building
Object-based communication

- **Users**:
  - Minimum: 2 → 1 instructor and 1 builder.
  - Number of builders can be increased.
  - Instructor role can be assumed by a confederate participant.
- **Task performance**:
  - Example: Ratio between the number of constructed figures and the time spent in constructing them.
- **Objects**:
  - VR settings in various labs → Shape to easily model them → Cubes
  - Size: freedom to make them bigger to mover around them.
- **Environment**: No particular restriction (easiness to model them in VR).

Task: Block building
Object-based communication

- **Questionnaires:**
  - Quality of experience
  - Quality of interaction
  - Presence, social presence, enjoyment
  - Cybersickness
  - Behaviour:
    - Audio recording
    - Exploration movements

Task: Block building

Object-based communication

Lab	Systems	Participants	Type	Research Question
UR3	VR	2	Blackbox	Impact of aumount of information (audio / video)
UCL	Ubiq	1 instructor 2 builders	Blackbox	Effects of collaboration style, roles and immersion (AR, VR).
LUT	MS Teams vs. Mozilla hubs	2	Blackbox	Systematic evaluation of network

Collaborative block building

Anna Ferrarotti, Marco Carli
anna.ferrarotti@uniroma3.it, marco.carli@uniroma3.it

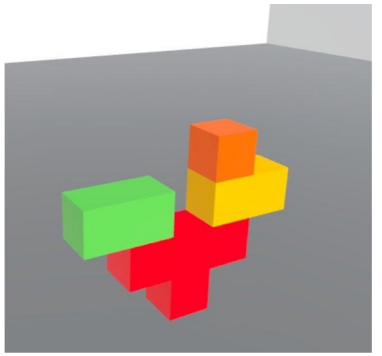
Roma Tre University

Test conditions

- The **remote user** (U1):
  - Has a representation of the target
  - Has a supporting app (for recording time)
  - Provides instructions to the builder in VR
- The **VR user** (U2):
  - Is in the virtual environment
  - Follows the instructions of the remote user
  - Builds the figures

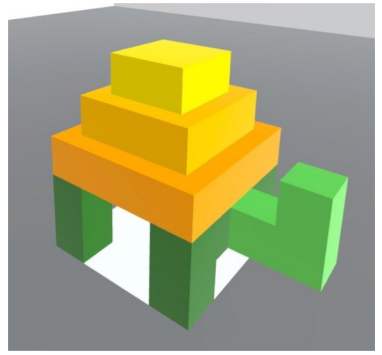
Conditions:

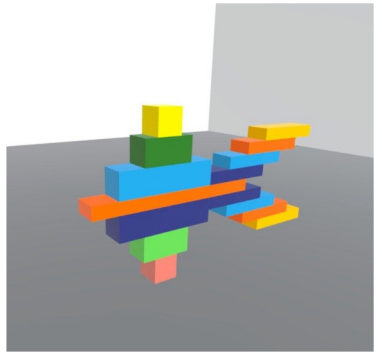
1. U1 and U2 see and hear each other
2. U1 sees and hears U2, U2 only hears U1
3. U2 sees and hears U1, U1 only hears U2
4. U1 and U2 only hear each other



Level 1 (5 pieces)

Level 2 (10 pieces)

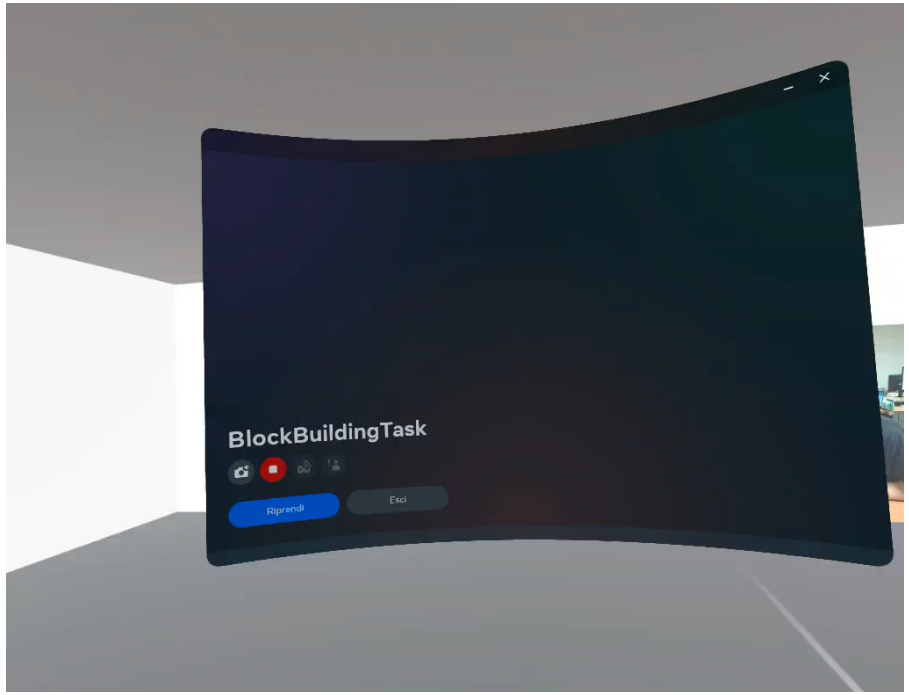




Level 3 (15 pieces)

Test setup

Remote user



VR user

Proposed measures

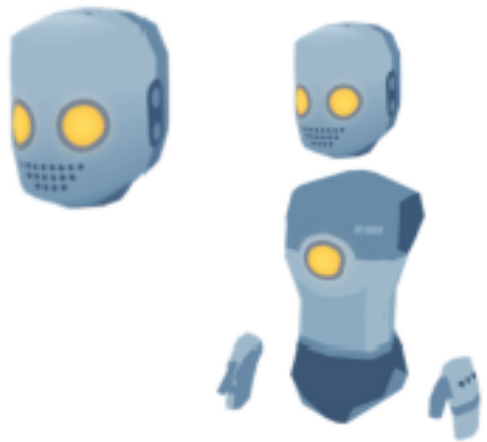
- Time per figure and total time
- Success rate (e.g., full success (structure solved in 2 minutes), success (structure solved between 2 and 3 minutes), borderline (between 3 and 4 minutes), and failed (more than 4 minutes)) - to be defined during pilot study
- Cybersickness scores
- Amount of exchanged traffic
- Movements
- Audio recording (e.g., recording the audio and segmenting the amount of time spent in interacting w.r.t. the total experiment time)
- Questionnaires (age, gender, previous experience, visual acuity, and color blindness tests, Quality of interaction, Social presence, Cybersickness, Quality of experience)

University College London

Experimental Protocol

Effects of level of ***immersion*** and ***embodiment*** on collaboration outcomes for a block building.

- Two users in VR vs two users in AR
- Co-located users in both immersive conditions
- Instructor with avatar face only or avatar full body
- 4 rounds per study
  - VR - face instructor
  - VR - full body instructor
  - AR- face instructor
  - AR - full body instructor
- 2 structures of 8 blocks to build each rounds
- Each participants experiments ent
  - Both embodiment conditions
  - Both immersive conditions

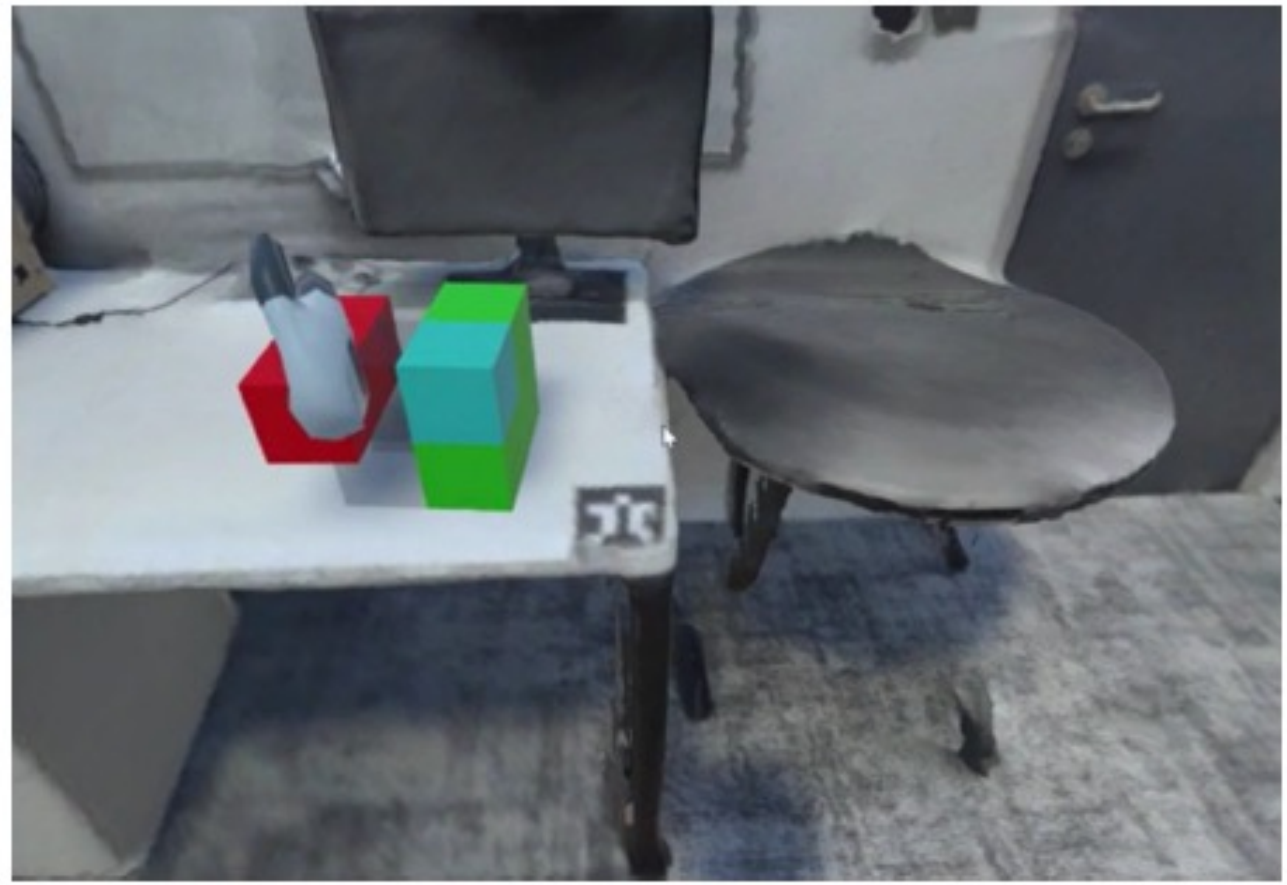


The Instructor
- See a representation of full structures and blocks
- Give the instructions

The Builder
- Only see blocks
- Follows the instructions and builds the figures

VR environment and building blocks

- Environment:
  - 3D Model Office and real office
- Objects:
  - Virtual cubes of same size
  - Virtual cubes of different colors
- Selection Technique:
  - Seek and Grab
- Building technique:
  - No gravity
  - Snap the cube to the target



Metrics

- Task completion time measured
- Speaking duration
- Behavioral measures logged
  - Head and body pos/rot
  - Body pos/rot
  - Look at object/instructor/builder/hands
  - Grab object

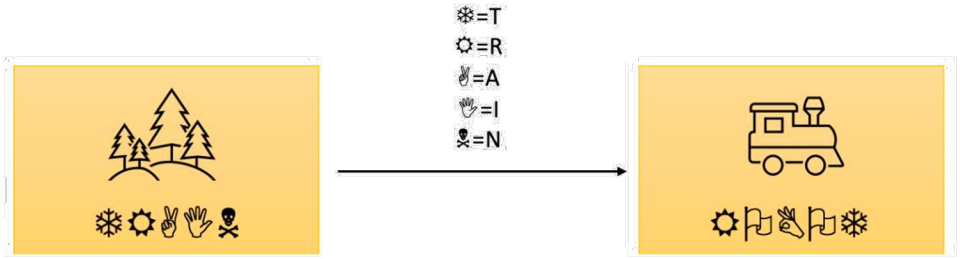
Questionnaires

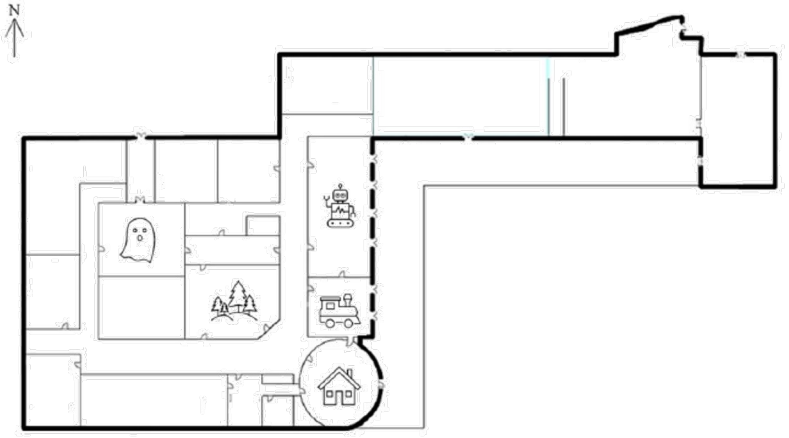
- Pre-study
  - Age, Gender, previous vr experience, color blindness
- Post-study
  - Presence, Social presence, Enjoyment

Task: Treasure Hunt

Environment-based communication

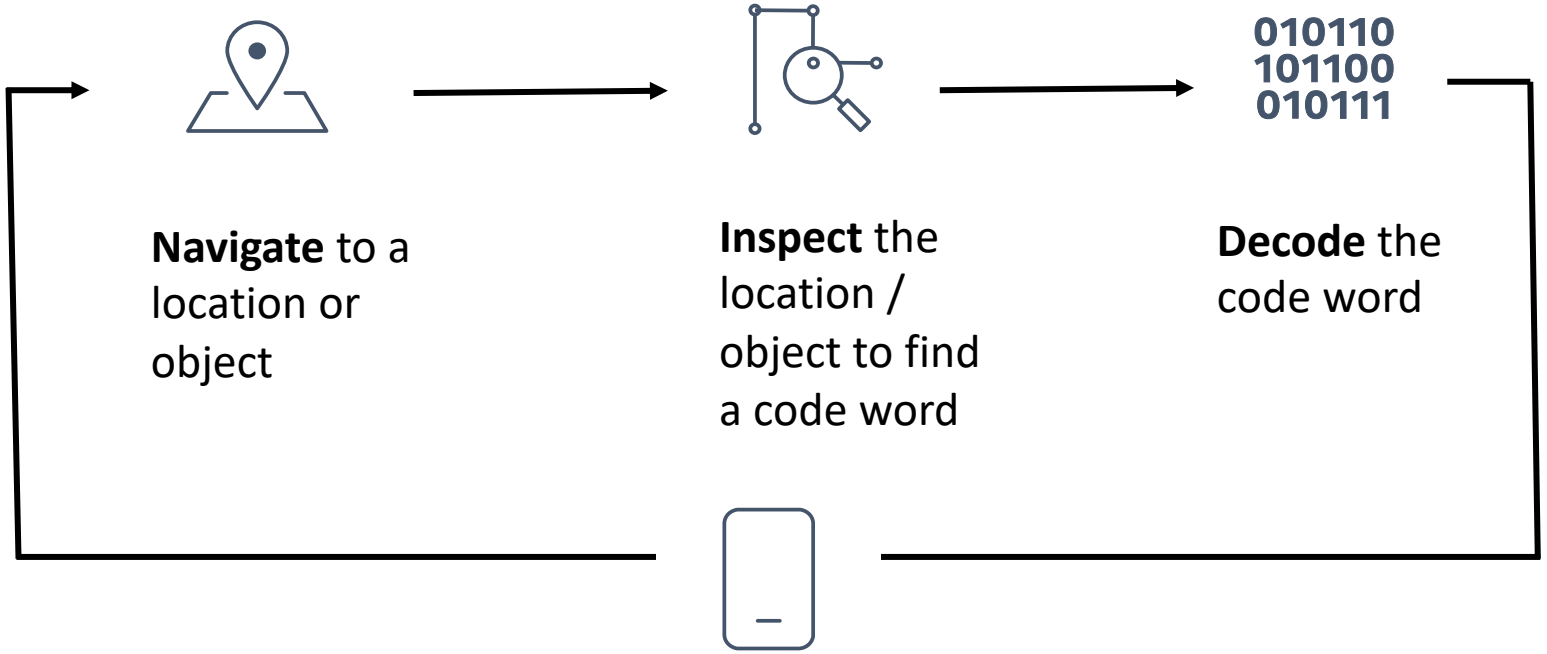
- Task focused on the **exploration** of the environment either in a **symmetric** (collaboration) or **asymmetric** (remote support / remote training) way.
  - Information needed to solve the task is divided between locals and remotes.
  - Include remote-only use case (teleoperation)
- **Treasure Hunt** task: systematically finding hidden objects or locations, each containing a clue leading to the next





Task: Treasure Hunt

Task Design

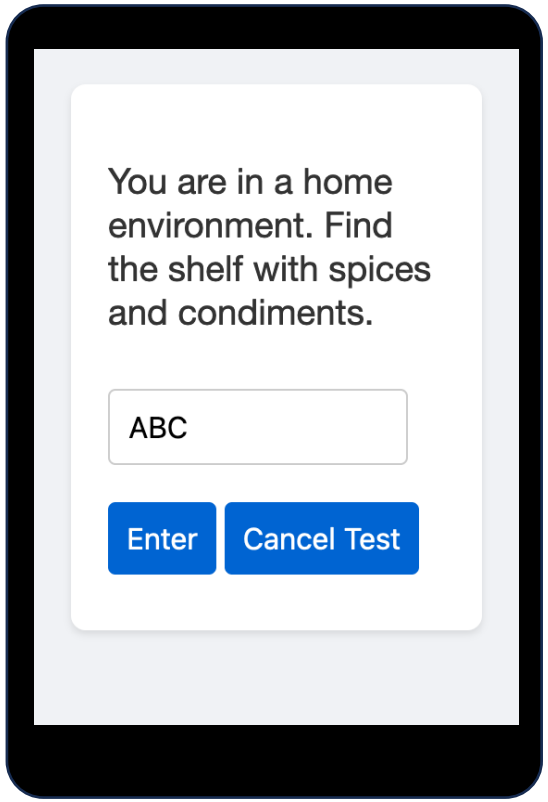


Navigate to a location or object

Inspect the location / object to find a code word

Decode the code word

Enter the code word in a web application

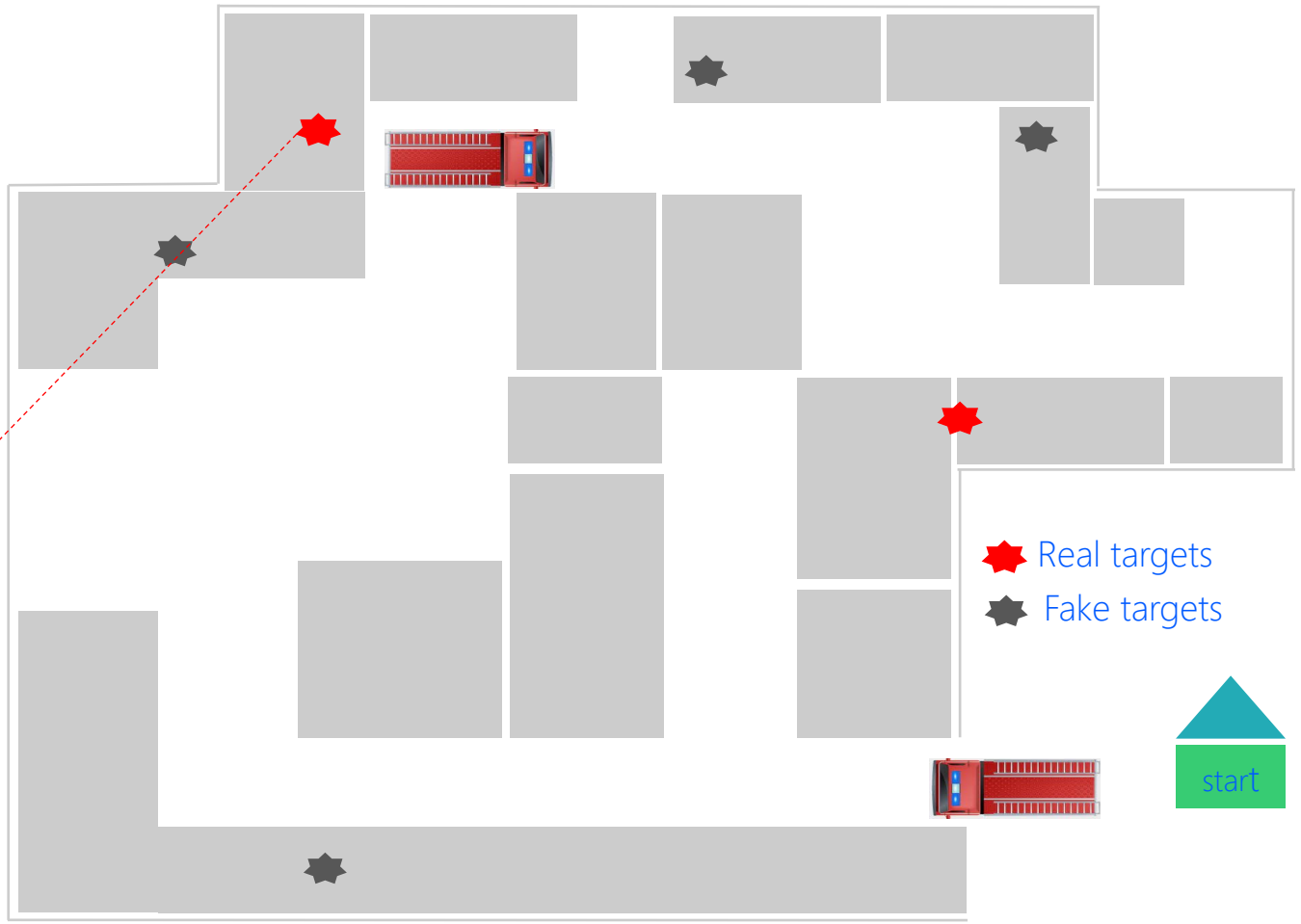


Sample application available at https://github.com/zerepolbap/vqeg-pixc

Task: Treasure Hunt

Navigation

1. User navigates to a target location
2. User inspects the location and retrieves the code for the next.
3. Users decodes de codeword and inputs in the application.
4. The next location is revealed.
5. GOTO 1



Task: Treasure Hunt

Navigation

Navigation: the user has to find a specific place in the room/area (e.g. a poster image)

a) Search: the user does not know the location
b) Guided navigation: a remote user / aid guides the navigation.
c) Simple: navigation is direct (e.g. object is visible).

Use cases:

- Remote support for exploration (e.g. remote expert guiding the field support).
- Tele-operated driving
- Remote first aid (the remote knows what to search, the local knows where it is)

Task: Treasure Hunt

Inspection

- **Inspect** an object and reveal N (3-4) code letters.
- There may be more active zones (code letters)
- Remote user has the list and order of zones to reveal.
- 2D objects (poster) o 3D objects (box) can be used

Use cases:

- Remote expert (expert knows what to explore)
- Local expert (student ask questions about the topic)

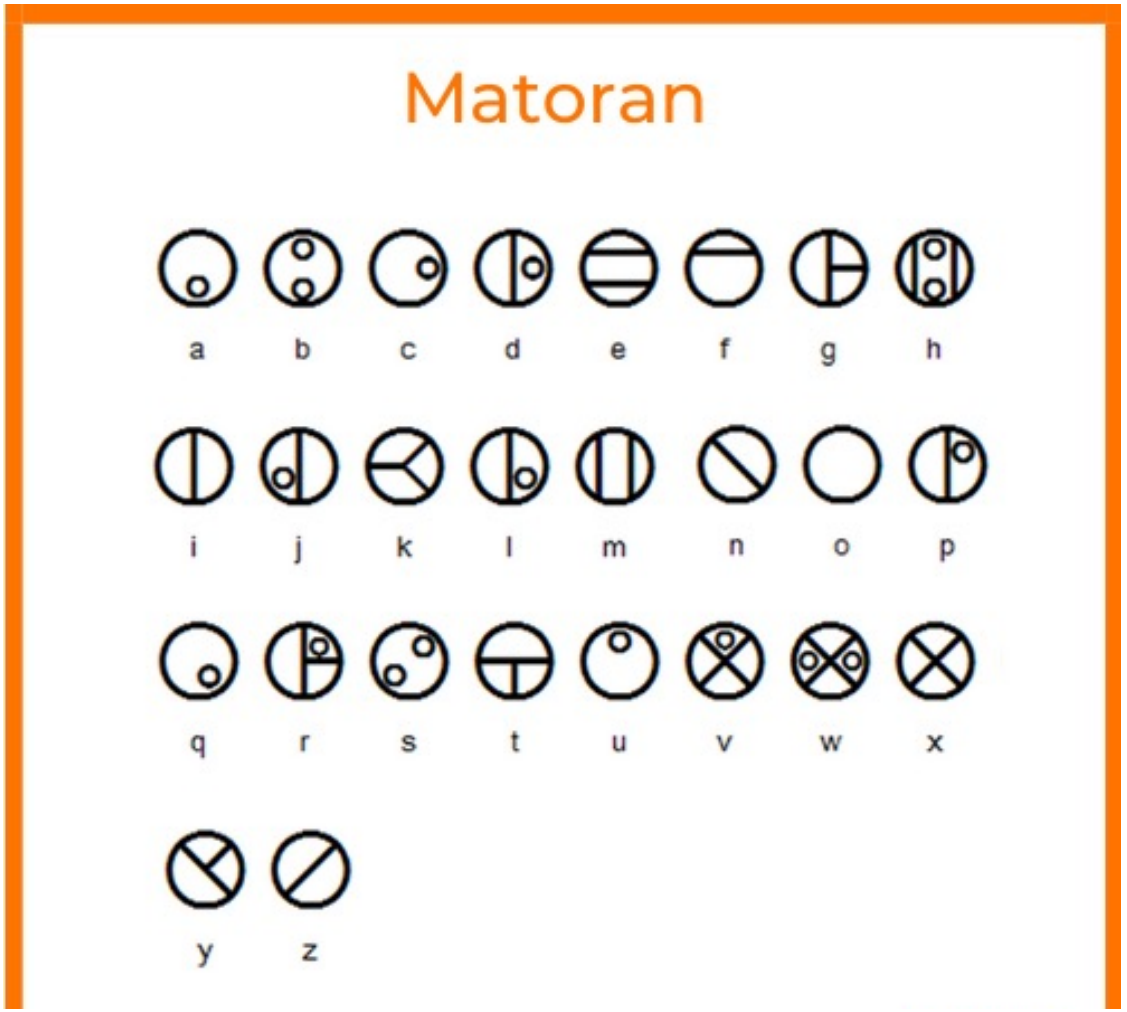
Task: Treasure Hunt
Decoding

- **Decode** a word using a simple substitution code.
- Code is visual: codewords can be either shown or described.
- Dictionary may be local, remote, split or shared.

Use cases:

- Dictionary represents asymetric knowledge (e.g. remote expert)

Todo: check copyright of selected alphabet



Task: Treasure Hunt

Decoding

- We have explored alternative codes based on human signs (to cover some health / communication use cases)
- **Discarded** as this is covered by Physiotherapy / Charade task

Task: Treasure Hunt

Task is very flexible!

Lab	Tecnology	Local	Remote	Navigation	Inspection	Decoding
Nokia	Immersive Telepresence	1	2	Simple	2D / 3D objects	Matoran code
Uni Padova	Shared Virtual Reality	1	1	Guided (remote user)	2D objects	Matoran code
RISE	Tele-operated Driving	0	1	Guided (navigation aids)	2D objects	TBD
ATLAS + Ericsson	Mobile Augmented Reality	1	1	Simple	3D objects	N/A

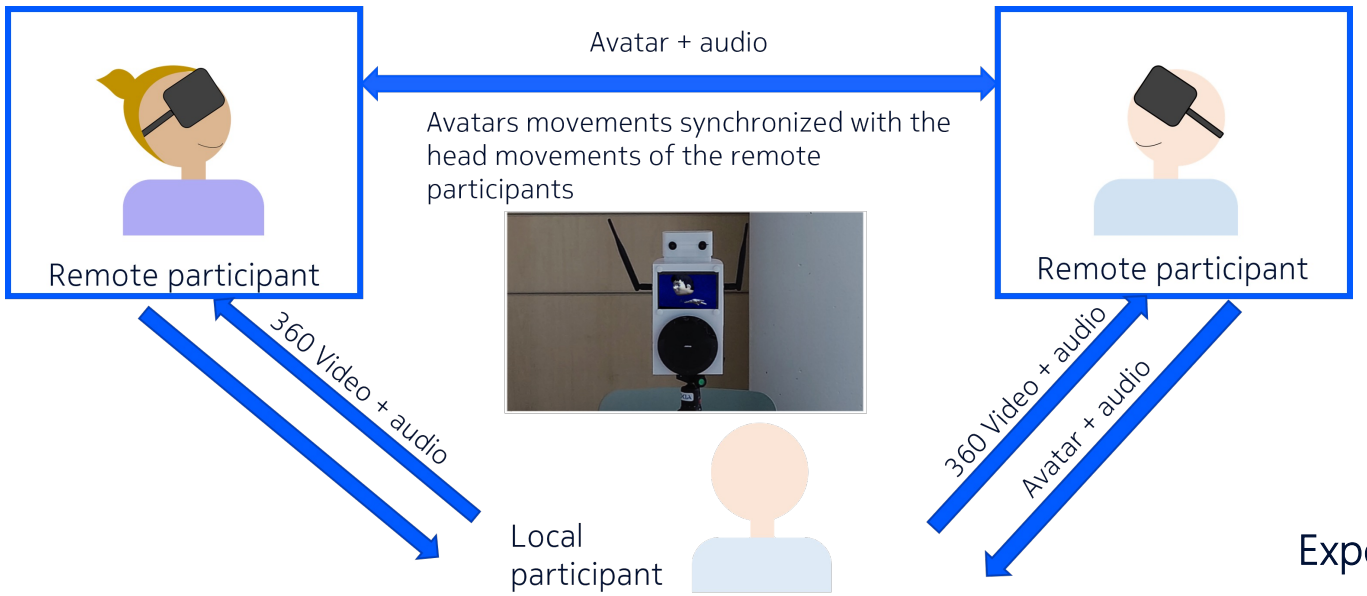
Treasure Hunt

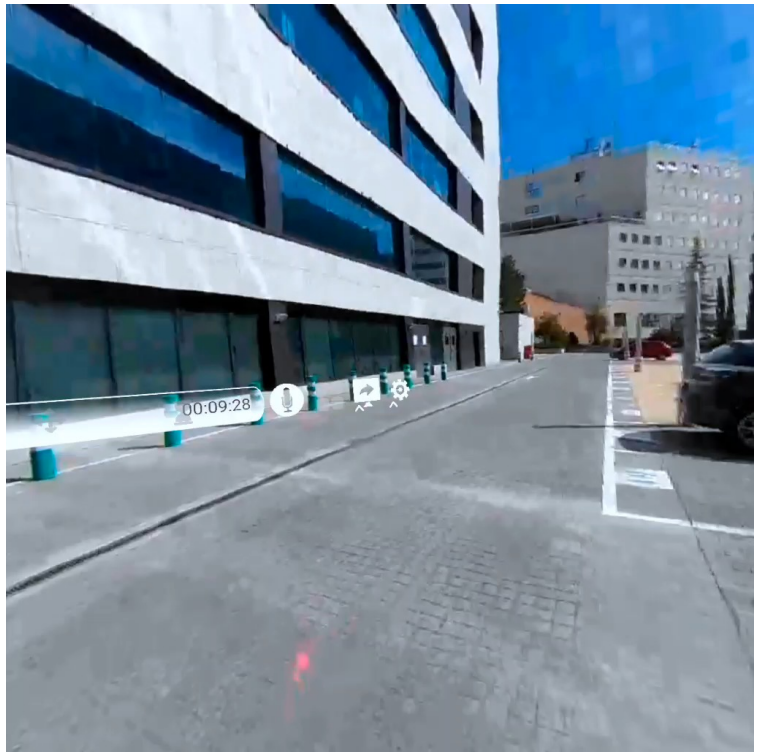
Marta Orduna
Mario Utiel
Pablo Pérez

  |

System

The Snowl





Navigation through 3 rooms,

each one associated with one use case:

- teleassistance
- guided tour
- remote technical support

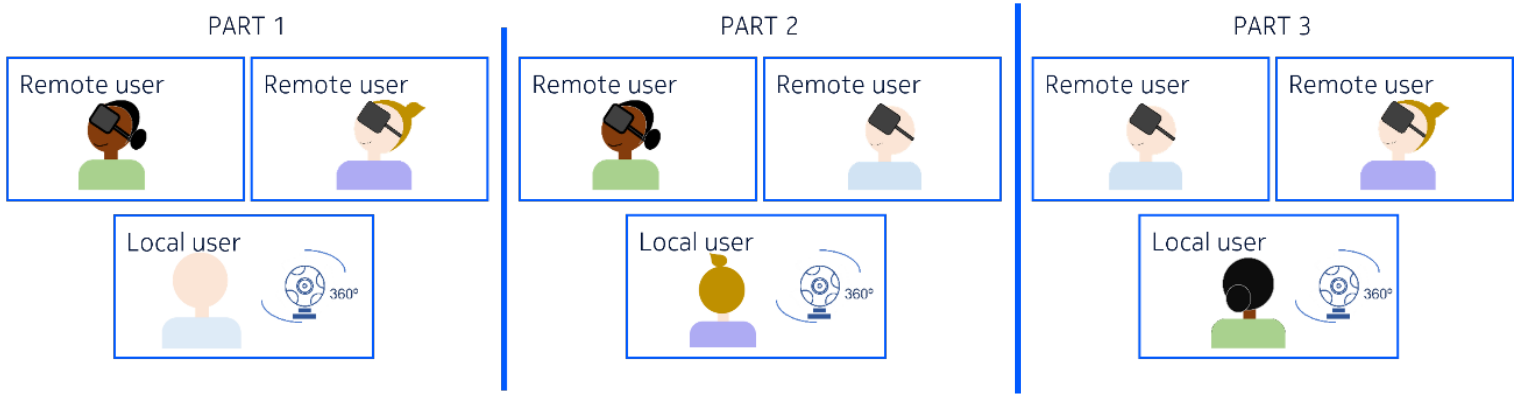
Experimental conditions:

- 360-degree Video with Bitrate: **8 Mbps**
- 360-degree Video with Bitrate: **8 Mbps and Magnifying glass**
- Conventional **2D Videoconferencing** (i.e., videocall with cell phone)

The experimental condition is maintained during one game: teleassistance + guided route + remote technical support

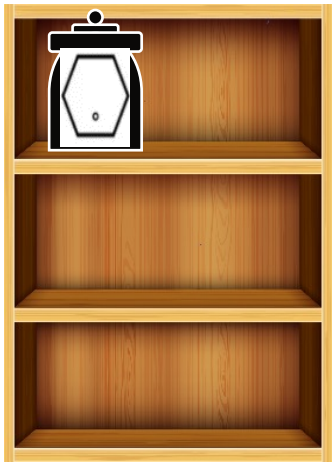
Test session

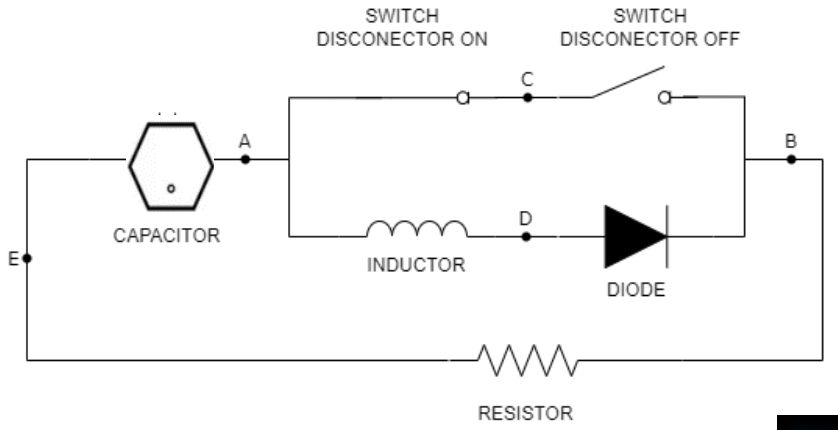
Part = 3 ROOMS (Use cases) x 2 OBJECTS/ROOM (Use case) x 3 Tasks



- In each part, one experimental condition is tested
- Remote users see the same information (matoran alphabet) in the virtual environment

TASKS:





Measures

Video:

- Please, rate the quality of the video

Audio:

- Please, rate the quality of the audio
- Audio recording

Latency:

- Did you perceive any reduction in your ability to interact during the communication due to delay?

QoE:

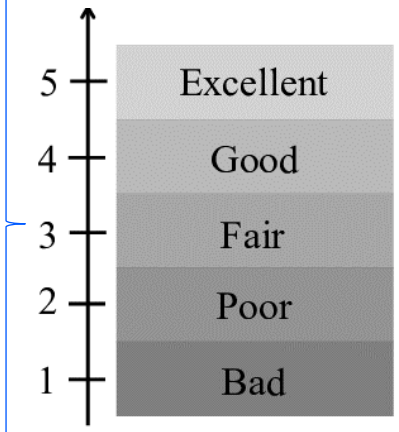
- Please, rate the quality of the experience

OWL:

- [Local] Track OWL movement
- [Remote] Trajectories

Task Performance:

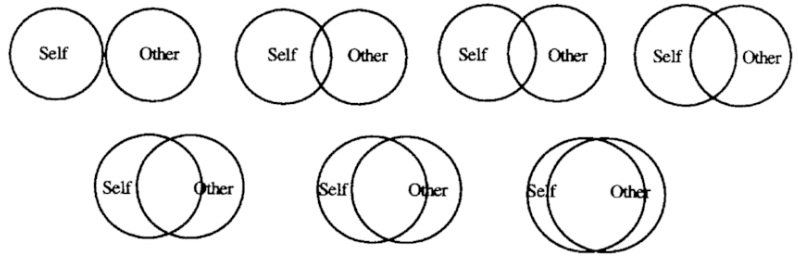
- Success rate and task success completion time



- Spatial presence (subsampled scale Witmer98, Pérez21)
- Mental Load (NASA-TLX)
- Simulator Sickness (VSR, P.919)
- Social presence, 7-level scale (Orduna22)
  - SP1: I felt that people were talking to me
  - SP2: I felt that I was listening to the others in the video
  - SP3: I felt I was present with the other people in the video
  - SP4: I felt like the people in the video could see me
  - SP5: I felt I was actually interacting with other people

Inclusion of Other in the Self (IOS) scale

Please circle the picture below which best describes your relationship



Collaborative treasure hunt

Sara Baldoni, Federica Battisti
sara.baldoni@unipd.it, federica.battisti@unipd.it

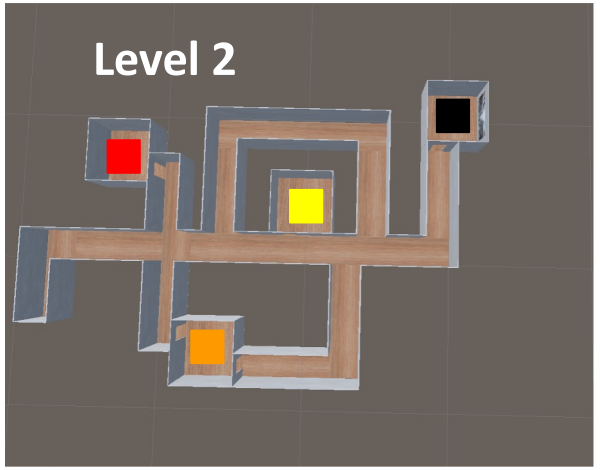
University of Padova

Test conditions

- The **remote user**:
  - Has a map of the maze
  - Has a dictionary
  - Has a supporting app (for providing hints and recording time)
  - Provides instructions and decodes the hints
- The **VR user**:
  - Is in the maze
  - Follows the instructions of the remote user
  - Finds the hints

Conditions:

1. The remote user sees where the VR user is, and there is audio communication
2. The remote user does NOT see where the VR user is and there is audio communication



Test setup

Remote user

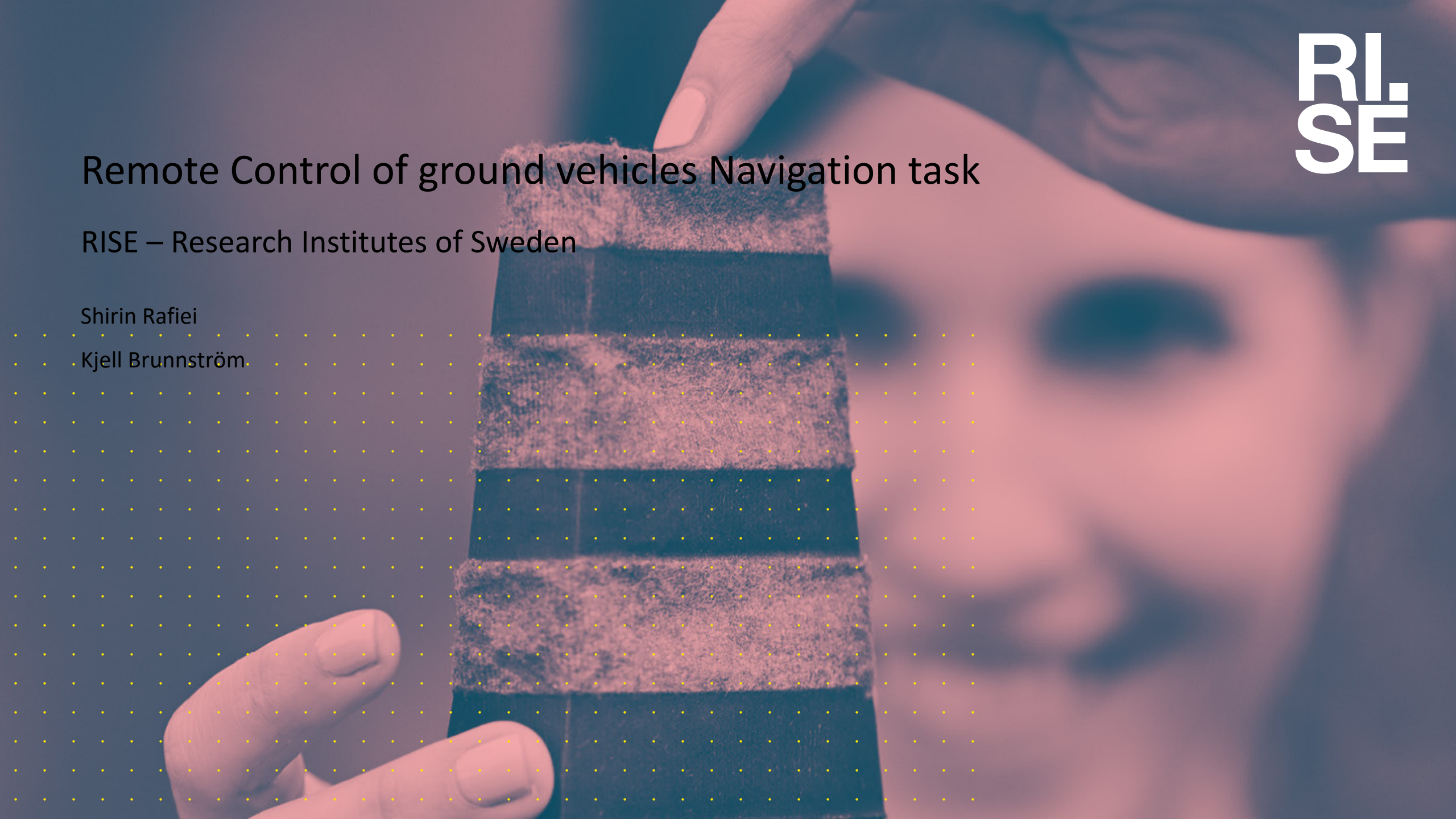
VR user

Scenarios:

- Museum
- Remote support (electrical support, computer support, VR application support)

Proposed measures

- Time per poster and total time
- Success rate (e.g., full success (task solved in 2 minutes), success (task solved in 3 minutes), borderline (between 3 and 4 minutes), and failed (more than 4 minutes)) - to be defined during pilot study
- Cybersickness scores
- Amount of exchanged traffic
- Movements
- Audio recording (e.g., recording the audio and segmenting the amount of time spent in interacting w.r.t. the total experiment time)
- Questionnaires (age, gender, previous experience, visual acuity, and color blindness tests, Quality of interaction, Social presence, Cybersickness, Quality of experience)



Remote Control of ground vehicles Navigation task

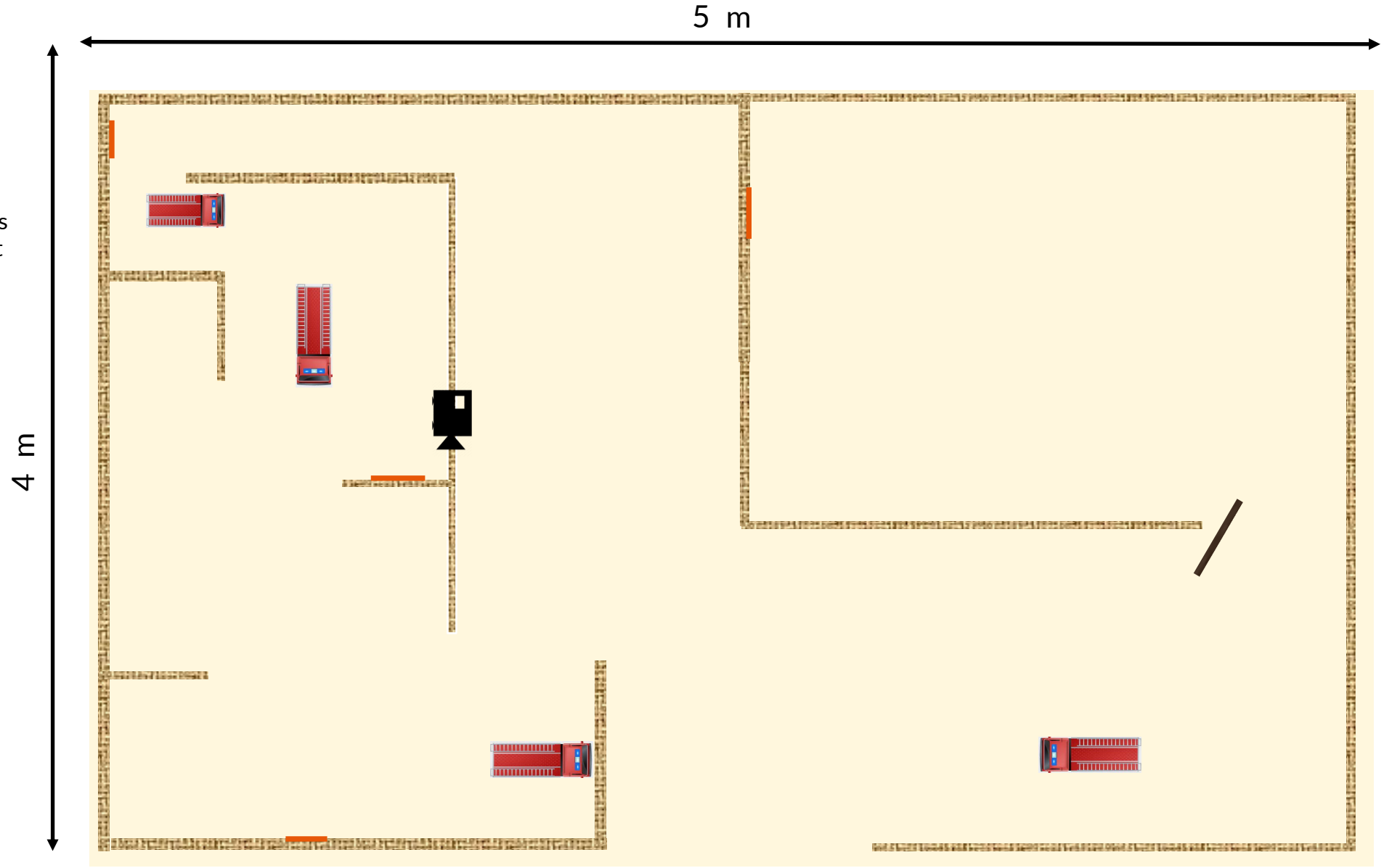
RISE – Research Institutes of Sweden

Shirin Rafiei

Kjell Brunnström

Test setup

1. User navigates to a target location
2. User inspects the location and retrieves the code for the next location.
3. Users go to the next location for a new inspection.



Test setup

Remote site

On site

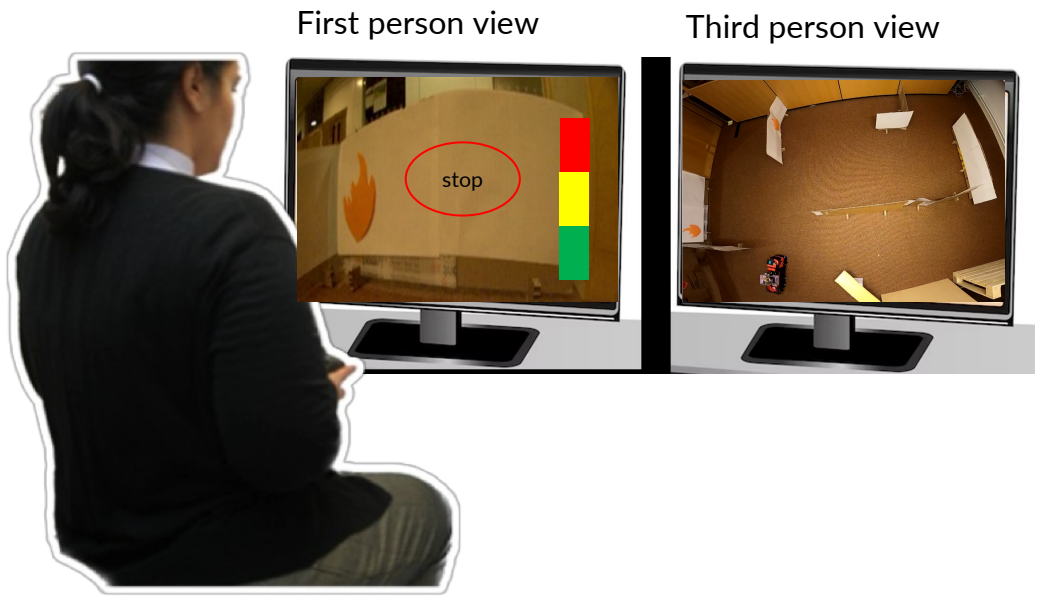
First person view

Third person view

Test conditions

We need to provide two views for one operator. Instead of using the second operator in the loop for assist, we will provide this support in finding the treasure hunt using augmentation (e.g, text, visualized depth aiding, …)

- We will add environment augmentation (on first person view) to support remote users in navigating and inspection tasks.



Test run	Combinations of the configurations		
	First Person view	Augmentations	**Third person view**
Trial 1	Yes	Yes	Yes
Trial 2	Yes	No	No
Trial 3	Yes	Yes	No
Trial 4	Yes	No	Yes
Trial 5	No	Yes	Yes
Trial 6	No	Yes	No
Trial 7	No	No	Yes
Trial 8	-	-	-

Proposed measures

Independent factors

- Augmented depth aiding
- Camera position 1_first person view
- Camera position 2_third person view

Dependent factors

- Objective Performance
  (Time for task accomplishment + Errors)
- Perceptual QoE
  (Depth perception and 3D understanding)
- Socio-emotional factors
  (Spatial presence, UX)
- Fatigue
  (Mental load and SSQ)

Shirin Rafiei

Shirin.Rafiei@ri.se

Kjell Brunnström

Kjell.brunnström@ri.se

Studying Mobile Spatial Communication

Comparing **Video Calls** and **Mobile AR**
For **Environmental Remote Communication**

December 2023

ATLAS-Ericsson research collaboration, 2023 - 2024

ATLAS

ERICSSON

Studying Mobile Spatial Communication

Status Quo: Mobile video calls are often used to **share spatial information about environments and people**.

Problem: Viewing this information as a **video limits spatial comprehension and interaction.**

Idea: A **more spatial and immersive** remote communication experience is possible **using Mobile AR**

Building on top of DualStream [Vanukuru et al. ISMAR 23]

Mobile Video Call

Spatial AR Call

Study Plan

Conditions:

- Mobile Video Call
- Mobile AR Spatial Call

  Custom prototype for both conditions

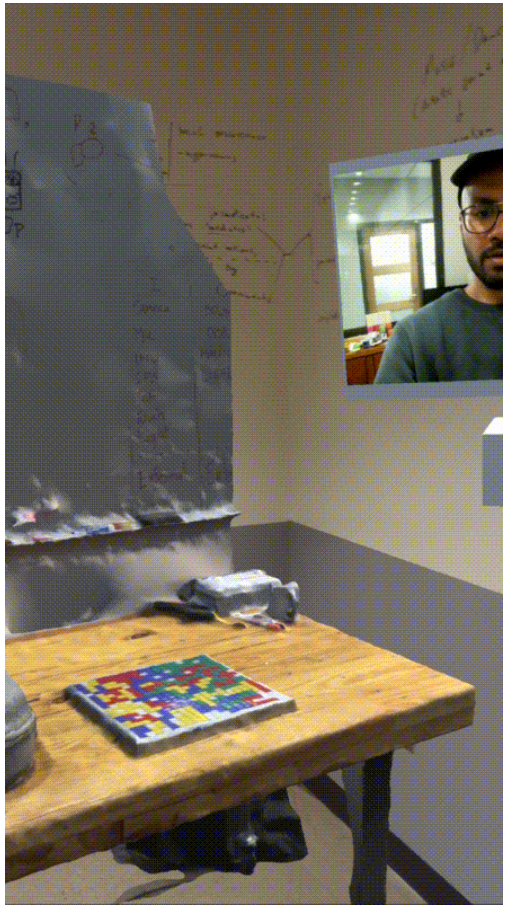
Task:

Participants will be given a **tour of a remote environment**, and will assist the researcher in **locating objects in the environment**.

Study Design:

Within-subjects, counterbalanced. For each condition, we will use a different remote environment (similar lab spaces, 10 * 10 sq. ft).

Mobile Video Call (Remote View)



Mobile Spatial Call (Remote View)

Measures

Questionnaires *(embedded in prototype)*

- **Social & Spatial Presence:** Subscales from Networked Minds [Biocca & Harms], Temple Presence Inventories [Lombard et al.]
- **Workload:** Subscales from NASA TLX [Hart]

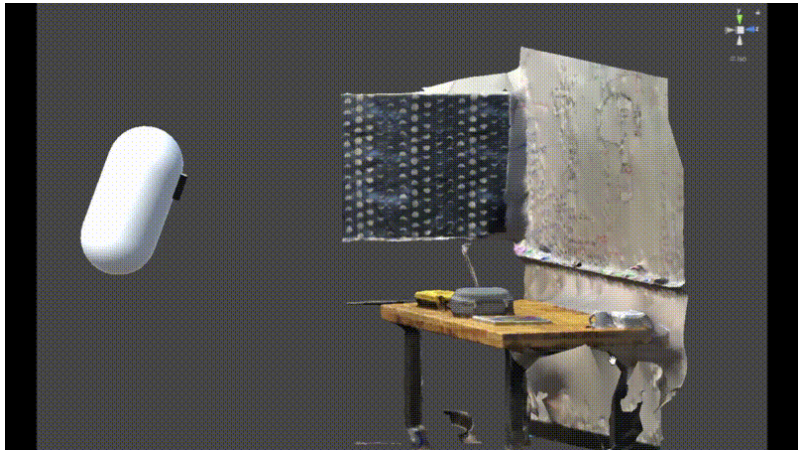
Screen, Video Recordings: For interaction analysis

In-app movement and interaction logs: For playback, quantitative analysis and visualization

Post-task interviews

Pilots: December 2023 - January 2024
Study: February 2024 - March 2024

In-app questionnaires



Spatial Playback from logs

Task: Treasure Hunt

Measurement highlights (1/4)

	Objective Performance		Behaviour	
	Task Performance	**Error / Precision**	**Audio**	**Motion**
Nokia	Task Success Completion Time	Success Rate	Conversation statistics	Local Motion Remote Poses
Uni Padova	Time per poster Total time	Success Rate	Conversation statistics	Motion
RISE	Time for task accomplishment	User's error		
ATLAS + Ericsson	Content Recall Spatial Recall Time for search task		Conversation statistics https://chronoviz.com/ ATLAS TI	In-app movement logs, Video

Common / similar objective metrics. Tools and statistical analysis to be discussed.

Task: Treasure Hunt

Measurement highlights (2/4)

	Video	Audio	Delay	Depth	QoE
Nokia	ACR	ACR	ACR*		ACR
Uni Padova					ACR
RISE				Define depth perception tasks and subjective rating Eyetracking?	ACR
ATLAS + Ericsson	Interview				

"Technical QoE" using ACR scale. Details on non-conventional factors (e.g. depth) TBD

Task: Treasure Hunt

Measurement highlights (3/4)

	Mental Load	Sim. Sickness
Nokia	NASA-TLX	VSR (P.919)
Uni Padova		SSQ
RISE	NASA-TLX	SSQ
ATLAS + Ericsson	NASA-TLX (subsampled)	

Physiological effects (Mental Load, Simulator Sickness) using standard questionnaries

Task: Treasure Hunt

Measurement highlights (4/4)

	Socio-emotional Factors			
	Spatial Presence	**Social Presence**	**QoI**	**UX**
Nokia	Subsampled PQ [Witmer 98, Perez 2021]	Social Presence [Orduna22] IOS [Aron92]		
Uni Padova	Subsampled PQ [Witmer 98]		QoI [Gupta16]	
RISE	IPQ			Mindmap evaluation method for: Interface helpfulness User satisfaction Usability
ATLAS + Ericsson	TPI (subscales)	NM		SUS

Diversity of socio-emotional measures. Further discussion to converge where needed

User feedback
Discussion

Objective Performance		Behaviour	
Task Performance	**Error / Precision**	**Audio**	**Motion**
Task Success Time (completion, total, …) Recall	Success Rate User's error	Conversation statistics	Local Motion Remote Poses In-app movement logs Video

Subjective performance - MOS				
Video	**Audio**	**Delay**	**Depth**	**QoE**
ACR Interview	ACR	ACR*	Define depth perception tasks and subjective rating Eyetracking?	ACR

Mental Load	**Sim. Sickness**
NASA-TLX	SSQ VSR (P.919)

User feedback
Discussion

Spatial Presence	Social Presence	QoI	UX
Subsampled PQ [Witmer 98, Perez 2021] IPQ TPI (subscales)	Social Presence [Orduna22] IOS [Aron92] NM	QoI [Gupta16]	Mindmap evaluation (Interface helpfulness) User satisfaction Usability SUS

Next steps

- Start writing ITU recommendation
  - Draft by April 2024 (ITU meeting).
- IMG meeting to discuss results:
  - Beginning of April 2024.
- Present final results:
  - Next VQEG plenary meeting in June/July 2024.
  - Joint paper on the whole test plan.
  - Exploitation of results of the individual tests by each lab.

Lab	Tests finished by
NOK	March 2024
CWI	March 2024
UPM	January 2024
AGH	March 2024?
PAD	March 2024?
UR3	March 2024?
TSU	February 2024
TUI	March 2024?
USU	March 2024
UCL	March 2024?
KUM	Spring 2024
RISE	March 2024?
ERI+UCB	March 2024

VQEG

IMG Test plan on Immersive communication systems

2023-12 – VQEG Plenary Meeting (Online)

Pablo Pérez (Nokia, Spain)
Jesús Gutiérrez (Universidad Politécnica de Madrid, Spain)
Kamil Koniuch (AGH University, Poland)
Ashutosh Singla (CWI, Netherlands)