VQEG

IMG Work Plan - what's next?

2022-05 – Rennes

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.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

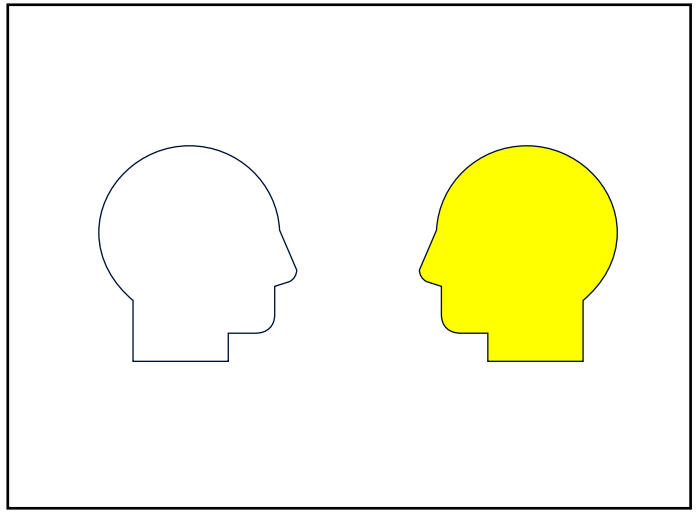
We did a bit of literature research since last meeting

- 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, arXiv preprint: https://arxiv.org/abs/2205.05953

Looking through a window

Project Starline - Google



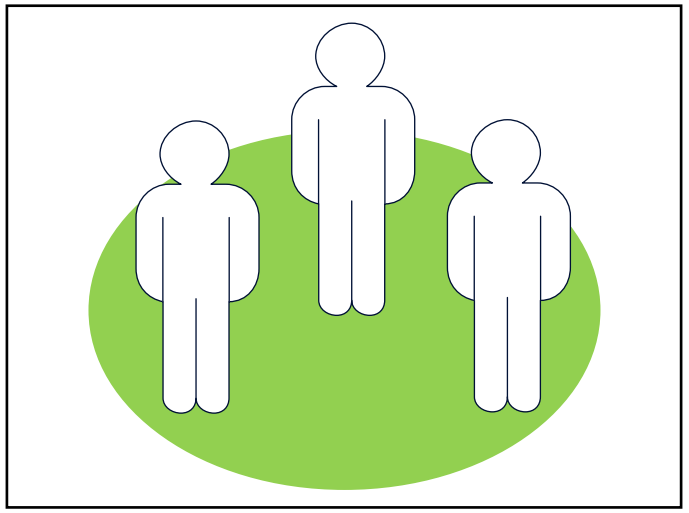
Visual communication

A Shared Virtual Environment

Social VR Platforms

AltspaceVR™

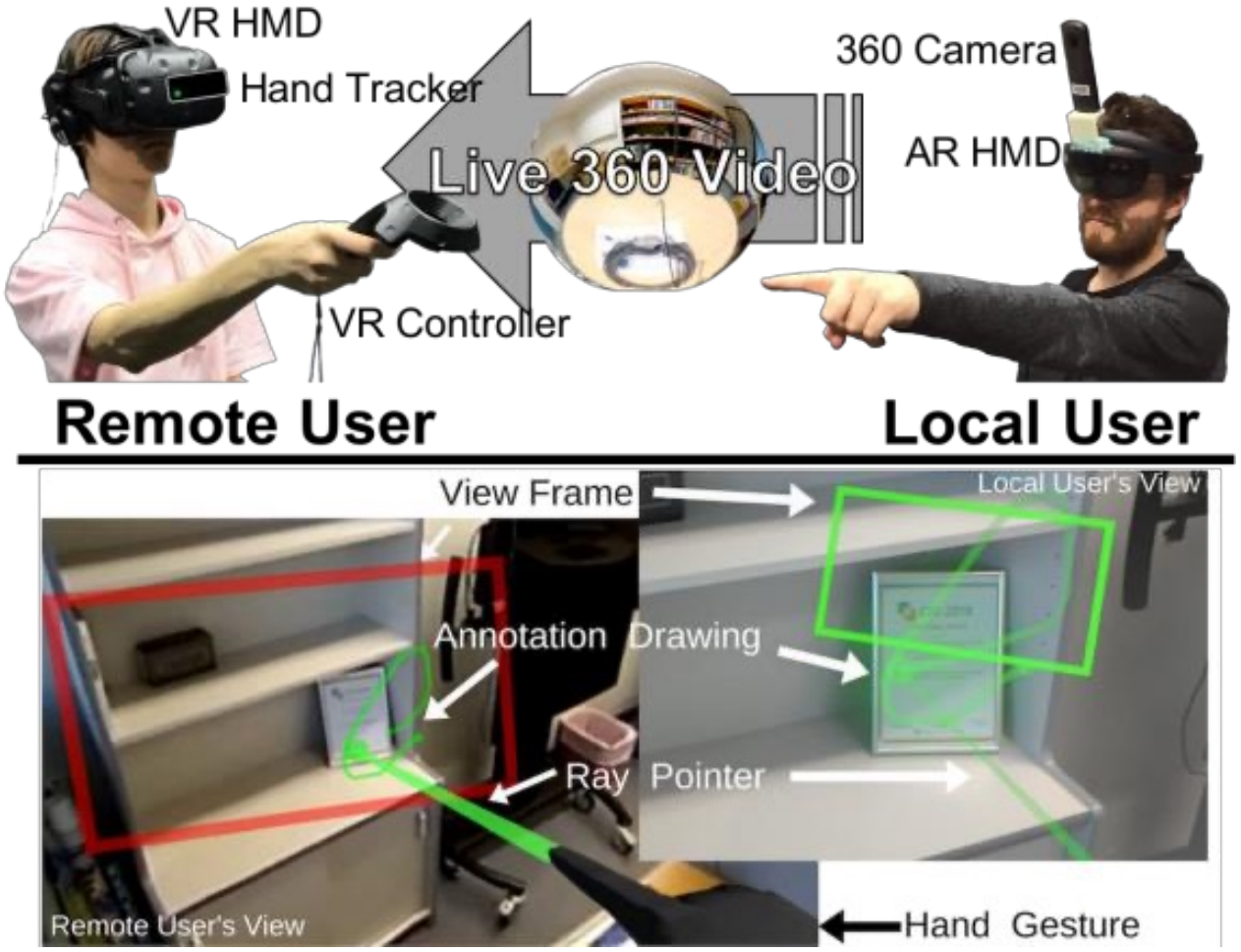
Shared immersion

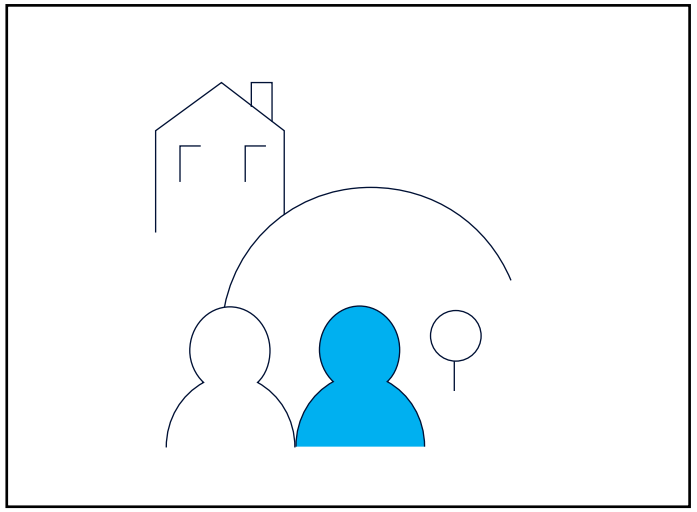


vTime® XR

Be Sociable

Remote Collaboration



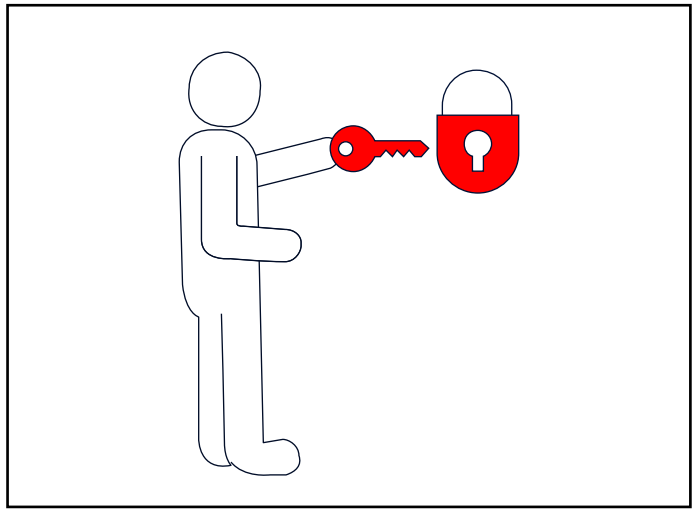


Remote Presence

Teo, Lee., Billinghurst, & Adcock (2019). *Investigating the use of different visual cues to improve social presence within a 360 mixed reality remote collaboration*.

Interaction with the World

Embodied Interaction



Brunnström, Kjell, et al. "Latency impact on Quality of Experience in a virtual reality simulator for remote control of machines." (2020):

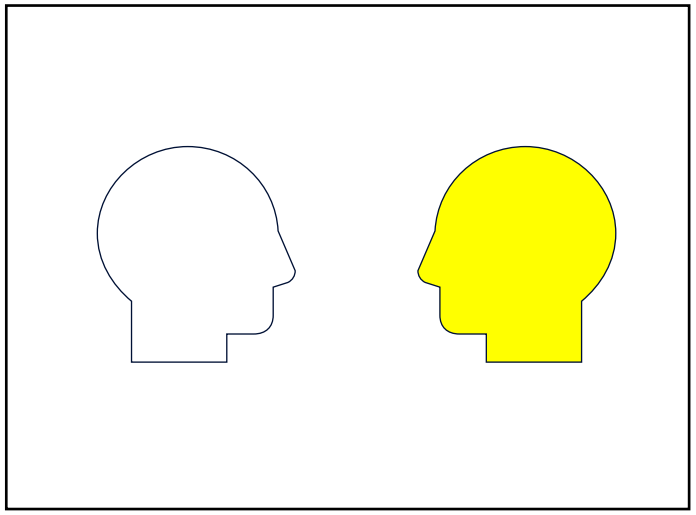
VQEG

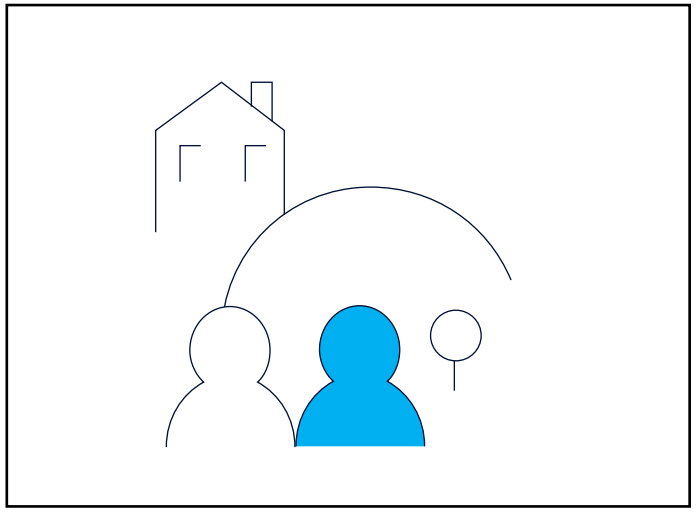
FACE

Visual Communication

Face to face

I see you





VISIT

Remote Presence

Shoulder-to-shoulder

I see what you see

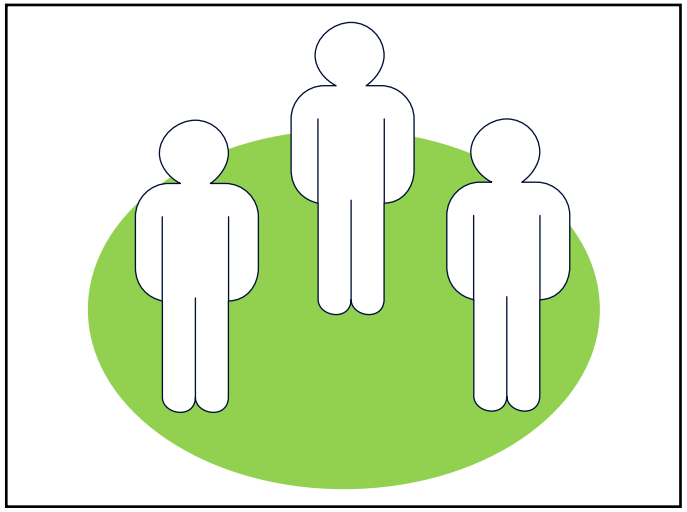
Technological / perceptual building blocks

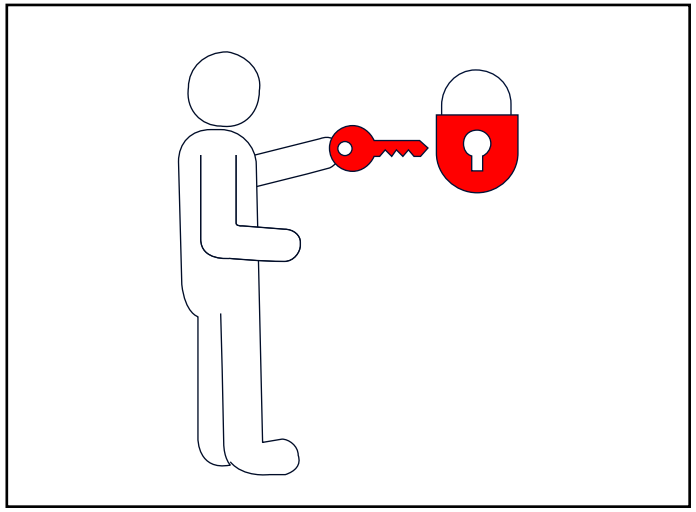
MEET

Shared Immersion

Hand-in-hand

I am with you





MOVE

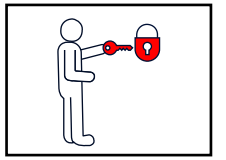
Embodied interaction

Hands-on

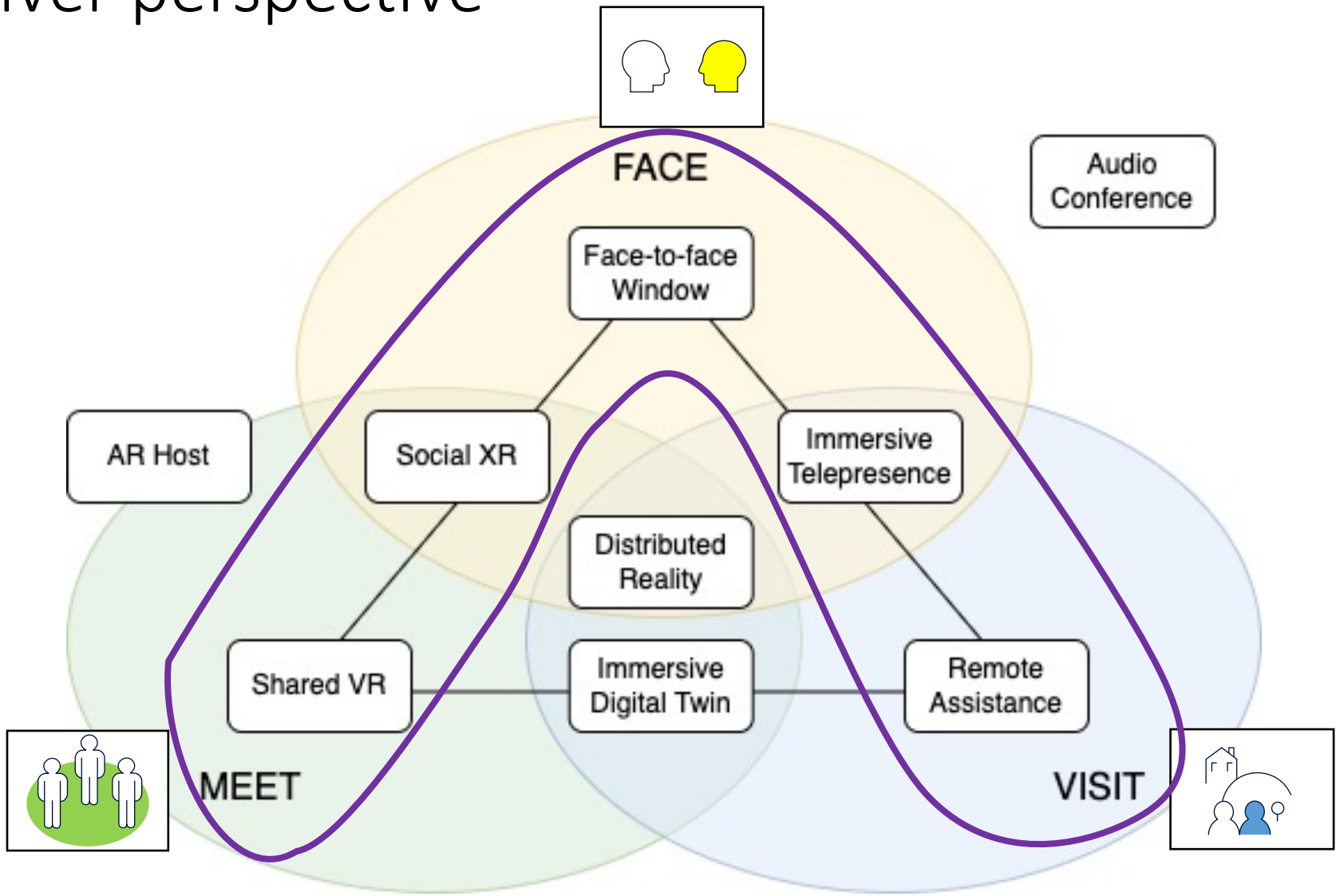
I control objects

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.

The basic questions

- Which is the minimum number of **tasks** which
  - allow us to test all the basic features, and
  - apply to as many systems as possible?
- How (if anyhow) are those tasks affected by **technical limitations**?
- Which are the **low-level QoE features** that we can measure?
  - Task performance
  - Visual quality, responsiveness, etc.
- Which are the **high-level QoE features** that we can measure?
  - Place/social presence
  - Are standard metrics possible (i.e. applying to all scenarios)
- How can we **measure**?
  - **Questionnaires** (subjective), **behavioural metrics** (e.g., gaze/head movements), physiological measures (e.g., EEG, EDA, ECG), and performance measures (e.g., time logging, success rates, etc.).

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	–

Overview of existing works

System **archetype** according to the previous classification.

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
2017)	Shared VR	Play games	–	–	Solving puzzles (pieces)	Embodiment types: no self-avatar, self-avatar, and face2face	IT for trust	CT for task performance
9)	Shared VR	Photo sharing	Comment on shared photos	–	–	Face2face, skype, and Facebook Spaces	SocialVR for social presence PMRI for emotions	–
. (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
018)	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
18)	Social XR	Play a game	–	–	Pong (two-player game)	System performance	Feedbak on presence and overall quality	–
2018)	Social XR	Conferencing	Negotiation	–	Assembly of lego blocks	Audio only, video fixed to HMD, and video fixed to host world	NMM-SPI	–
1)	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	–

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. (2019)	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
	XR	Play a game	–	–	Pong (two-player game)	System performance	Feedbak on presence and overall quality	–
	XR	Conferencing	Negotiation	–	Assembly of lego blocks	Audio only, video fixed to HMD, and video fixed to host world	NMM-SPI	–
	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
	nce	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
	nce	Remote exploration	–	Explore remote environment	–	Three ways of interaction	IPQ for presence NMM-SPI for social presence and emotions SSQ for cybersickness	–
	nce	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	–

Context of use:
- Conferencing
- Play games
- Watch/share videos/images
- Remote collaboration
- …

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,	RGQ: visual and audio quality
Prins et al. (2018)	Social XR	Play a game	–	–	Pong (two-player game)	System perf		
Lawrence et al. (2018)	Social XR	Conferencing	Negotiation	–	Assembly of lego blocks	Audio only, vi and video fix		
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. with game c		
Lee et al. (2018)	Remote Assistance	Remote collaboratio	–	Find a set of target objects in the task space	Place the target objects on the desk	Dependent independe		
Young et al. (2019)	Remote Assistance	Remote exploration	–	Explore remote environment	–	Three wa interact		
Teo et al. (2019a)	Remote Assistance	Remote collaboratio	–	Identify objects	Decorate a bookshelf placing objects	No cues, hand ges and hand gestur		
Teo et al. (2020)	Remote Assistance	Remote collaboratio	–	Find a set of target objects in the task space	Place the target objects to a specific location	360 image r 360 projection mode	SMEQ for workload SEQ + 3 custom questions for global QoE	CT for task performance
Wang et al. (2020)	Remote Assistance	Remote collaboratio 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 collaboratio	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 collaboratio	–	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	–

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

Overview of existing works

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

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
					–	System performance	RGQ: Social presence, interaction, explorability, and global QoE	RGQ: visual and audio quality
					Pong ayer game	System performance	Feedbak on presence and overall quality	–
					of lego blocks	Audio only, video fixed to HMD, and video fixed to host world	NMM-SPI	–
					h environment ck 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
					he target 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
					–	Three ways of interaction	IPQ for presence NMM-SPI for social presence and emotions SSQ for cybersickness	–
					a bookshelf g 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
					target objects ific 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
					mbly of 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
					mbly of 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
					–	System performance	–	Semi-structured interview for task performance
					–	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	–

Conditions:

- Analyzed system vs. previous/alternative technologies and or F2F
- Display types
- Visualization of pointers/cues (remote collaboration)
- Ways of interaction between host and visitor.
- Embodiment types
- Positioning of the visitor representation
- Latency
- Use of controllers
- Overall performance of the system

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
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 spatial for presence NMM-SPI for social presence SMEQ for workload SSQ for cybersickness	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	IPQ for presence NMM-SPI for social presence and emotions SSQ for cybersickness	–
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 SMEQ for workload	SUS for usabilty
Anjos et al. (2019)	AR Host	Play games.	Solve riddles	–	–	System performance	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
Kasahara et al. (2017)	Immersive Telepresence	Cleaning up a lab room	–	Locate objects and clean them	–	Video Stabilization	TQ: co-presence, interactivity, and explorability NASA-TLX for mental workload	TQ: Visual and audio quality
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 NASA-TLX for mental workload	SUS for usabilty CT for task performance
Zhang et al. (2020)	Immersive Telepresence	Telepresence	–	Locate / indicate remote user's gaze	–	Distance to avatar and display (AR, tablet)	–	Semi-structured interview 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	SSQ for cybersickness	HM for non-verbal behaviours
Brunnström et al. (2020)	No communication	Remote control	–	–	Control a crane to load logs	Latency	MEC for spatial presence NMM-SPI for social presence SSQ for cybersickness SMEQ for workload	SEQ for task performance
Pérez et al. (2021)	No communication	Escape room game	–	–	Manipulate game objects	Real hands vs. VR controllers	NMM-SPI for social presence SMEQ for workload	SEQ for task performance
							RC: Comfort, immersive, and overall quality	RC: Picture, responsiveness, and task accomplishment quality
							WS for presence Embodiment DREQ for Global QoE	–

Features:

- **High-level features**: focus on the user (cognitive or psychological constructs)
- **Low-level features**: evaluate technical aspects of the system

Overview of existing works

Features:

- High level: mainly through questionnaires
  - Presence (sometimes evaluated in terms of spatial, social, and co-presence)
  - Workload
  - Simulator/cyber-sickness
  - Quality of Interaction
  - Interactivity
  - (Social) Connectedness
  - Emotions
  - Explorability
  - Attentiveness and reaction/gauging
  - Embodiement
  - Immersion
  - Comfort
  - Trust
  - Global QoE
- Low level: questionnaires and objective measures
  - Task performance
  - Visual quality
  - Audio quality
  - Usability
  - Exploration
  - Responsiveness
  - Latency
  - …

Overview of existing works

Questionnaires:

Questionnaire	Measure, N. items, Scale, Subscales, Factors
Interpersonal Trust (IT) Johnson-George and Swap (1982)	Interpersonal trust in social situations. 21 items for male version, 13 items for female version. 9-point Likert scale. Subscales: reliableness (male & women), emotional trust (male & women), and general trust (male).
Nasa Task Load Index (NASA-TLX) Hart and Staveland (1988)	Workload of task. 6 items 21-point Likert scale Measuring mental, physical and temporal demand, performance, effort and frustration.
Simulator Sickness Questionnaire (SSQ) Kennedy et al. (1993)	Users' levels of cybersickness symptoms. 16 items. 4-point scale. 3 subscales: Nausea (N),Oculomotor (O), and Disorientation (D)
Subjective Mental Effort Question (SMEQ) Zijlstra (1993)	One question that measures the mental effort. 1 item. 9 labels scale (from "Not at all hard to do" to "Tremendously hard to do").
Witmer and Singer (WS) Witmer and Singer (1998)	Sense of presence. 32 items. 7-point Likert scale. Subscales: Involvement/Control, Natural, Auditory, Haptic, Resolution, Interface Quality Factors: Involvement/Control, Natural, Auditory, Haptic, Resolution.
System Usability Scale (SUS) Brooke (1996)	Usability. 10 items. 5-point Likert scale.
Igroup presence questionnaire (IPQ) Schubert et al. (2001)	Presence. 14 items. – Subscales: spatial presence, involvement, experienced realism.
Networked Minds Measure Social Presence Inventory (NMM-SPI) Harms and Biocca (2004)	Social presence and emotions. 36 items. 9-point Likert scale. Subscales: co-presence, attention allocation, perceived message understanding, perceived affective understanding, perceived emotional independence, and perceived behavioral independence.
MEC Spatial Presence Questionnaire (MEC-SPQ) Vorderer et al. (2004)	Spatial presence. 32, 48, 64, items (4, 6 and 8 items per each of the 8 subscales), 5-point Likert scale. Subscales: attention allocation, higher cognitive involvement, suspension of disbelief, spatial situation model, spatial presence self location, spatial presence possible actions, domain specific interest, visual spatial imagery.
Single Ease Question (SEQ) Sauro and Dumas (2009)	Assessment of how difficult users find a task. 1 item. 7-point Likert scale.
Pictorial Mood Reporting Instrument (PMRI) Vastenburg et al. (2011)	Emotions. 9 items. 5-point scale. Nine moods: excited, cheerful, relaxed, calm, bored, sad, irritated, tense and neutral.
Embodiment Gonzalez-Franco and Peck (2018)	User embodiment on immersive experiences. 25 items. 7-point Likert scale. Subscales: body ownership, tactile sensations, location of the body, external appearance, and response to external stimuli.
Feedback Prins et al. (2018)	Requirements gathering to understand user expectations for social VR. 6 items. 7-point Likert scale. Subscales: Social presence, interaction, explorability, visual quality, audio quality, and overall quality.
Requirements Gathering Questionnaire (RGQ) Gunkel et al. (2018)	Short questionnaire to gather feedback form users' immersive experiences. 2 items. 5-point Likert scale. Subscales: Presence and overall quality.
SocialVR Questionnaire Li et al. (2019)	Social and interactive experiences in immersive media. 24 items. 5-point Likert scale. 3 subscales: Presence / Immersion (PI), Social Meaning (SM), and Quality of Interaction (QoI).
Distributed Reality Experience Questionnaire (DREQ) Perez et al. (2019)	Presence and quality aspects. 10 items. 5-point scale. Subscales: presence, video quality, cybersickness and quality of experience.
Remote Control (RC) Brunnström et al. (2020)	QoE aspects. 6 items. 5-point Likert scale Subscales: picture quality, comfort quality, immersive quality, overall quality, responsiveness quality, and task accomplishment quality.
Tele-collaboration Quality (TQ) Wang et al. (2020)	Social presence (slightly modified from Gupta et al. (2016) and Harms and Biocca (2004) to better reflect the experiment). 7 items 7-point Likert scale. Subscales: co-presence, interactivity, explorability, visual quality, audio quality, fatigue.
Visual Quality of Experience (VQoE) Li et al. (2021)	Visual quality of self and others' volumetric representations. 2 items. 5-point Likert scale.
Holoportation Questionnaire (HOLO) Lawrence et al. (2021)	Global QoE for holoportation systems. 7 items. 5-point scale for 6 items and 7-point scale for one. Subscales: presence, attentiveness, personal connection, reaction-gauging, engagement, closeness, eye-contact.

Next steps

- Identify interested people/groups:
  - UPM (Spain), Nokia Bell-Labs (Spain), CWI (The Netherlands), RISE (Sweden), TU Ilmenau (Germany), Ghent University (Belgium), Wuhan University (China), University of Surrey (United Kingdom), AGH University of Science and Technology (Poland), University of Brasilia, Meta (US).
  - **You?**

- Available systems to be evaluated:
  - CWI SocialVR, Nokia Owl, UPM FVV Live, RISE Crane control
  - **More?**

Next steps (II)

- Then:

1. Characterize the available systems (according to the archetypes)
2. Determine the system factors to be tested:
   - The system is being compared with others: verify whether all the systems under consideration share the same fundamental elements or not.
   - Different configurations/conditions of the same system are being tested.
3. Identity a set of tasks
4. Identify the relevant QoE factors to measure
5. Refine the tasks so that they cover the QoE factors
6. Select the right assessment tool for each QoE factor:
   - Questionnaires, sensors… as non-intrusive as possible.

- To be done in future audio calls…