25/05/2022

Virtual visits:
a video-meeting tool to enhance personal communication

Marina Álvarez Mérida
Máster en diseño de interacción y experiencia de usuario

TRABAJO FINAL DE MÁSTER
Tutor: R. Ignacio Madrid López
PRA: Ferrán Giménez Prado y Enric Mor Pera

TFM realizado en colaboración con **TNO** (Netherlands Organisation of Applied Scientific Research).

Supervisor: Dr. Alexander Toet

FICHA DEL TRABAJO FINAL

Título del trabajo:	Virtual visits: a video-meeting tool to enhance personal communication
Nombre del autor:	Marina Álvarez Mérida
Nombre del consultor/a:	R. Ignacio Madrid López
Nombre del PRA:	Enric Mor Pera
Fecha de entrega:	05/2022
Titulación:	Máster en diseño de interacción y experiencia de usuario
Idioma del trabajo:	Inglés
Palabras clave	Diseño de interacción; experiencia de usuario; video comunicación; realidad aumentada

Resumen del Trabajo:

Nuestra forma de comunicarnos ha cambiado y cada vez hacemos más uso de herramientas de videoconferencia (VC) como Teams o Zoom. Como resultado de las restricciones sociales introducidas durante la pandemia de COVID-19, la popularidad de esta forma de comunicación se ha disparado recientemente. Aunque estas herramientas son una forma efectiva de mantener relaciones de forma remota, todavía están limitadas en cuanto a crear un sentido de conexión con las personas en el otro lado y la interacción no resulta tan natural como un encuentro físico.

El objetivo de este proyecto es presentar un nuevo sistema de videocomunicación, adaptado a las necesidades del usuario en materia de comunicación personal, que proporcione una experiencia más natural, fluida e inmersiva, a la vez que ofrece una usabilidad óptima.

Se ha seguido una metodología de DCU (diseño centrado en el usuario). Los resultados obtenidos de la investigación documental, investigación sobre las

necesidades de los usuarios y exploración sobre la aplicación de la realidad extendida en este contexto (en colaboración con TNO) han sido aplicados en la elaboración de un prototipo funcional. El diseño de la nueva herramienta de VC se basa en un enfoque multipantalla-multiplataforma y engloba la implementación de realidad aumentada. El prototipo ha sido evaluado positivamente y mejorado en base a los resultados de una prueba de usabilidad.

Abstract:

Social communication is increasingly performed via video-conferencing (VC) tools, such as Teams or Zoom. As a result of social restrictions introduced during the COVID-19 pandemic, the popularity of this form of communication has recently boosted. Although these tools are an effective way to maintain relationships remotely, they are still limited in creating a sense of social connectedness and the interaction does not feel as natural as a physical encounter.

The goal of this project is to present a new domestic video-communication system, adapted to user's needs on personal communication, that provides a more natural, fluid and immersive experience, while offering an optimal usability.

A UCD (user centered design) methodology has been followed. Insights obtained from desk research, research on user needs and exploration about the application of extended reality on this context (in collaboration with TNO) have led to a high-fidelity, and ultimately an interactive, prototype. The design of the new VC tool is based on a multiscreen-multiplatform approach and involves the implementation of AR. The prototype has been positively evaluated and improved based on the results of a usability test.

Table of contents

List of Figures

List of tables

Introduction

Humans are social animals; we have an intrinsic need of feeling connected to other people and a key element of that connection is communication. That is why communication technology has evolved so much in the last 100 years, but while this evolution has significantly increased the quantity of connections we can maintain, it hasn't been so successful preserving their quality in comparison with face-to-face interactions. (Apostolopoulos et al., 2012).

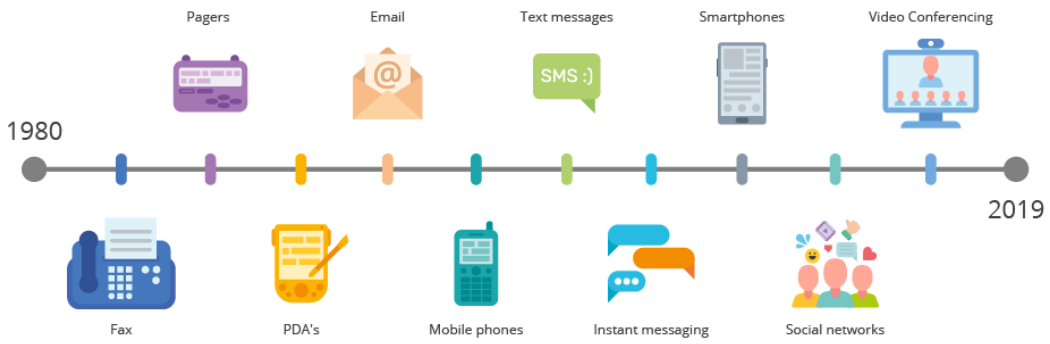


Figure 1. Evolution of communication technologies (trueconf.com)

1.1 Context and justification

Especially during a crisis like COVID-19 pandemic, that forces people to physical distancing and isolation, a sense of connection with others is most needed. Previous research on how mediated communication can help with that matter, suggests that virtual social experiences can provide a sense of togetherness (Miller, 2020). However, current communication technology is still limited, due to the lack of shared context and poor support of non-verbal cues, like eye contact or body language, that provide us with a great part of the meaning and emotional connection (Hauber,

2005). All of this contributes to the so called “zoom fatigue”: exhaustion experienced from using any kind of video calling interface to have virtual interactions with others (Sklar, 2020), that lots of people are experiencing lately due to a more intensive use.

In an attempt to provide videoconferencing systems with a “human touch” and improve remote social connection, the concept of Telepresence was introduced. It refers to a range of technologies that “allow a user appear to be present, feel like they are present or have some effect in a space the person does not physically inhabit” (Techopedia, 2017). Telepresence can include video teleconferencing tools, where a picture and audio stream are conveyed to a remote location, as well as more immersive tools and robotic installations that can actually help a user to accomplish tasks from a remote location.

Some famous examples are Cisco Telepresence systems or telepresence robots for remote collaboration, like the ones from AVA robotics.

Figure 2. Cisco IX5000

However, these systems are still mainly designed for and implemented in work-related environments and, common videocall systems for private use are still limited in providing a sense of closeness. New technologies like 3D displays, virtual reality (VR) and augmented reality (AR) start to be considered for such systems to improve their usability by providing a more immersive and realistic experience. In this line, the Dutch independent research organization TNO is creating an AR communication tool, based on previous work on social VR (Gunkel et al., 2019), that will be part of this project.

The main application of our system, would be daily communication with family and friends living apart from each other but, a successful system in this context, would be useful as well in other areas of application like:

- Teaching
- Communication and monitoring of people in healthcare institutions and elderly homes.
- Participation in remote activities like debates, workshops or the practice of sports from home
- Remote support (home-appliances, small offices)

1.2 Goals

Main Goal: To present a new domestic video-communication system, adapted to user's needs on personal communication, that provides a more immersive experience while offering an optimal usability.

Sub Goals:

1. Explore user's needs on remote personal communication in the domestic environment.
2. Explore trends and other immersive videoconferencing/telepresence solutions currently available.
3. Discover whether TNO's AR-based communication tool provides users with a better experience than a regular videocall system. More specifically, determine if it can it provide a more intense sense of social presence.
4. Define the characteristics of this system and develop a prototype that can be presented to stakeholders and evaluated by users.

1.3 Method

The approach followed in this project is UCD (user centered design), iterating through the 4 phases of the *double diamond* process:

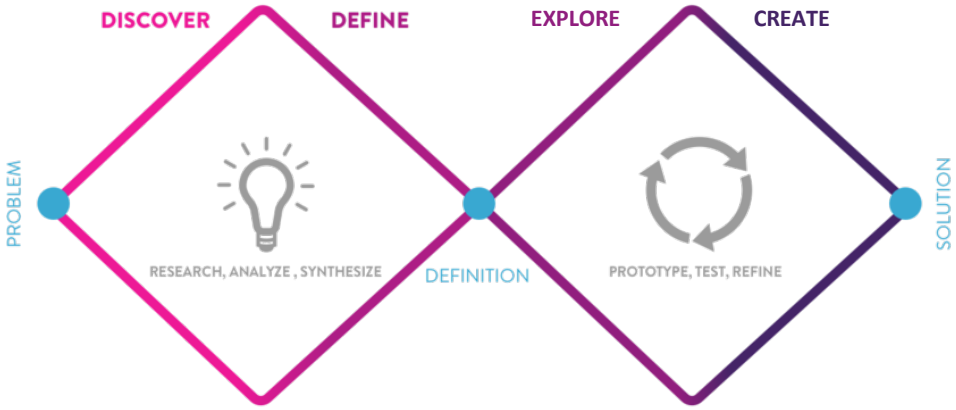


Figure 3. Double diamond design process

Note that, since the topic of this work is quite new, it would benefit from a thorough research phase. Therefore, the first diamond has more weight on this project.

1.4 Tasks and Planification

Research

Desk research: Bibliographic review of previous studies and trends related to telepresence and immersive communication, using different data bases and websites.

16 hours

Benchmarking: Research and comparison of current solutions, commercialized and prototypes, based on their characteristics. Results are showed in a comparative table.

10 hours

Research on user's needs: Interviews and questionnaire to people that live far from some family members/friends/partner and use video-calling tools to communicate with them frequently. Quantitative and qualitative data analysis: descriptive statistics and affinity diagram.

70 hours

Definition

Based on the results obtained in the research phase, Personas and User Journeys are developed to define user profiles and use cases, uncovering details of the context and possible pain points. These techniques help defining requisites that have to be implemented in the prototype.

18 hours

Exploration

Proof of concept (analysis and insights). The independent research organization TNO is currently performing a proof-of-concept experiment with users, on a prototype of the AR-based communication system they are developing, in which I am taking part. This section will contain the description of the experiment, analysis of results and main insights obtained about the use of augmented reality for this purpose.

30 hours

Creation/Prototyping

This section covers the design of the interaction (modality and definition) as well as the prototyping of the system. It contains tree of content for the app, user flow diagrams, sketches, hi-fi prototype, interactive prototype and a Storyboard showing the interaction with the system.

82 hours

Evaluation

Techniques used were heuristic evaluation and usability test with users (including a brief online whiteboard session to gather feedback).

26 hours

Improvements

Findings from the evaluation are applied on a new design iteration to improve the prototype.

12 hours

Documentation tasks

Corrections and layout of the report.

30 hours

Elaboration of video presentation

12 hours

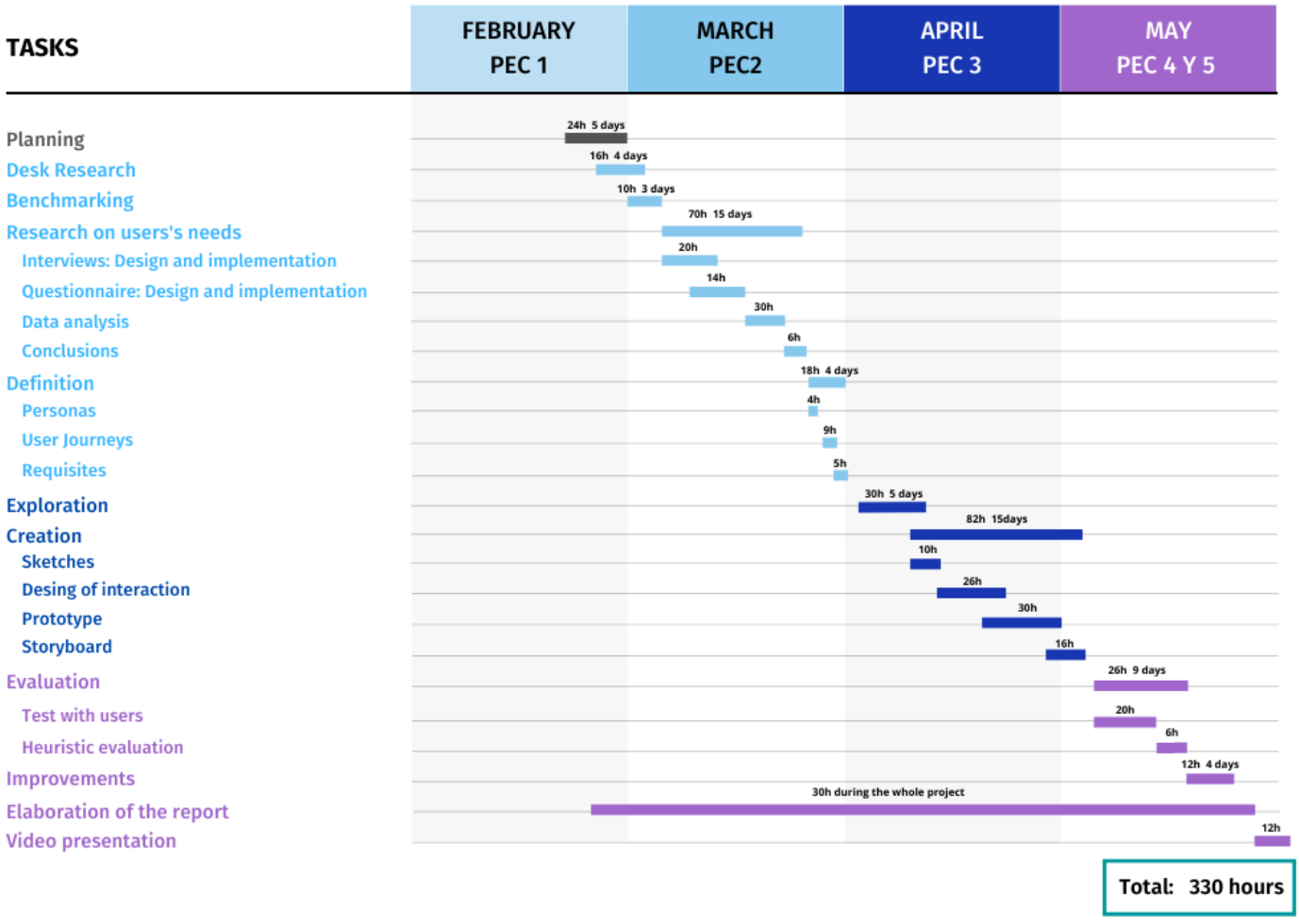


FIGURE 4. Gantt's diagram

1.5 Short description of other chapters

There will be 9 more chapters describing: the 5 phases of the process 1 last section with conclusions and recommendations, Glossary of terms, Bibliography and Annex.

2. **Research**
3. **Definition**
4. **Exploration**
5. **Creation**
6. **Evaluation**
7. **Conclusion & future work**
8. **Glossary**
9. **Bibliography**
10. **Annex**

Each chapter includes sub-sections covering the different techniques and products obtained.

Research

2.1 Desk research

This section contains a bibliographic review of previous studies and trends related to telepresence and immersive communication, using different data bases and websites, as well as a benchmarking of exiting products in this area.

When individuals are in need of social connection, and face to face is not an option, they tend to prioritize interpersonal media channels (such as video chatting or phone calling) that afford greater social presence and intimate communication. During Covid pandemic, interpersonal media use for social connection had a positive effect on users' psychological well-being by decreasing loneliness and increasing satisfaction with life (Choi & Choung, 2021).
Kirk et al. (2010) argue that we should focus on specific scenarios related to closeness with our loved ones, such as "communication between a grand-parent and a child that can't yet talk, the desire to partake in routines of the home when away or the desire to demonstrate that you have taken the time to focus on a person", as inspiration for future design.

2.1.1 State of the art

Advances in human-computer interaction (HCI) systems, including video mediated communication (VMC), are providing new opportunities to stay connected, especially to those who are separated by long geographic distances. But, compared with real face-to-face conversation, communicating through conventional videoconferencing tools, still feels artificial.

TRENDS

New functionalities are being introduced on classic videoconferencing tools, especially since COVID pandemic started, due to an increase in their use and new needs generated from this situation. *Together mode* from Microsoft Teams is the perfect example of trying people to feel closer to each other during an online meeting. It eliminates traditional squared grids putting all participant's video feeds onto the same virtual background (a classroom, a conference hall, etc), looking like they are actually sitting in a room together. You can see all members of the meeting at a glance.

The current availability of head-mounted displays (HMD) for consumers makes **virtual reality (VR)** a possible alternative to audio communication and 2D video conferencing. With VR technology, people are able to "meet" in a shared, immersive virtual environment and interact with virtual representations of each other. Such environments with multiple users are called *Social VR*. This technology has shown better results in terms of presence/immersion than a regular Skype call (Li et al. 2019) but, it is still limited to an avatar-based representation of the users and not a photorealistic image. As it can be expected, user representations that are not perceived as real deliver lower social presence impression (Yu et al. 2021).

Augmented reality (AR) allows for co-existence of real and virtual elements and do not replace the user's surroundings, as in VR. Instead, AR systems add 3D-registered virtual objects to the real world, which the user can typically interact with in real-time. It has been applied to mediated communication in different ways, such as remote assistance platforms, based on the use of AR glasses, to improve collaboration in manufacturing environments, or enhanced personal calls that allow you to add filters on your image or the image of the other person (like Snapchat or Facebook messenger). However, AR solutions for personal communication, oriented to provide a face-to-face, are more limited.

Figure 5. Remote assistance using AR glasses. (xrmeet.io)

The three-dimensionality of this technology has a big potential to achieve this, but the most effective way of applying it is yet to be defined.

3D / holographic displays have been explored for a while but, to the best of my knowledge, there are not commercialized communication systems based on this technology at the moment. This is due to its complexity. First of all, because its success is dependent on principles of human vision, like stereo disparity, motion parallax and accommodation. Static computer-generated holograms are capable of reproducing these cues, but this cannot be easily translated to images in real-time motion. There are three fundamental challenges: the computation of the 3D information, the transmission of the data from where it is captured to where it needs to be displayed (networks), and the reproduction of the holographic pattern on the 3D display (Blanche, 2021). However, Google is making progresses on this area with their Starline project.

There is much that can be done to develop video-communication systems to provide added value for the users, without discarding the video itself but strengthening its use (Kirk et al. 2010).

2.2 Benchmarking

Performing a benchmarking analysis has facilitated the discovery of products in the area of telepresence and immersive videoconferencing, digging into their functionalities and characteristics (based on some stablished criteria), to uncover advantages and limitations.

There are quite a lot of new immersive solutions to improve the perception of social presence on streaming content (like concerts or talks) but for this comparison I have selected just products that allow to have a real time two-way interaction. Also, solutions from different categories are compared, since one of the goals is to discover which technologies might serve best to our propose. Therefore, a representative product from each category/technology has been selected. The analysis cannot go very deep on specific functionalities and usability details tough, since the systems are fundamentally different and some are not commercialized yet.

Cisco Webex room panorama: premium video collaboration solution providing immersive video experience, rich content sharing, and co-creation experience for executive meeting rooms. All-in-one solution for medium-to-large rooms (webex.com).

Ohmni telepresence robot: home robot that allows people to be a part of event or project, monitor healthcare patients or connect with their families. OhmniLabs lets you take a step

further from regular videoconferencing by pushing your voice and image through a drivable rolling robot that allows to have a physical presence in a remote environment. It also has a tilting neck that amplifies the field of view.

Microsoft Teams (together mode): (previously explained) New Microsoft Teams feature that aims to connect all meeting attendants in one shared space. It uses AI mapping to cut out your face and shoulders, then creates an avatar for you, that is placed into a virtual environment alongside with the other participants.

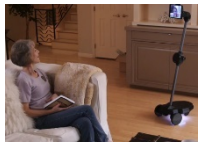
AltspaceVR: Social VR platform that consists of user-generated spaces called "worlds", which can be visited by other users, where virtual events are frequently held. In addition to these events, AltspaceVR is a social platform where individuals can gather, talk, collaborate, and be co-present in small to large groups. It is compatible with many head mounted devices and the interaction with commands takes place via floating menu as in most VR apps.

Matsuko app: It uses a combination of AR and AI technology to project life-like 3D holograms over your real-world environment using an iPhone and compatible XR devices. You simply aim the iPhone's camera at your face and start talking. Those with a compatible AR device like Nreal or Microsoft HoloLens 2, can view the iPhone-captured hologram in 3D real-time.

Table 1. Benchmarking comparative table

Name	**Cisco Webex Room Panorama**	**Ohmni**	**AltspaceVR**	**MS Teams. Together mode**	**Matsuko**
Stage	Launched 2020	Launched 2017	1st release 2015 (acquired by Microsoft)	Launched in 2020	In development
Type of product	Telepresence video conferencing system	Care telepresence robot	Social VR platform	Feature in a videocall app	Mixed reality conferencing app (human projection)
Context	Workplace: Team collaboration and executive meetings	Home-care	Group gatherings	Team meetings work/academic	Corporate environments (one-to-one)
Technology	4K display, 5K cameras, HD directional audio, Wall structure	High quality videocall + responsive tilting neck + drivable robot	Virtual reality	AI	Augmented reality (human projection)
Hardware requirements	Wall screen structure, HD cameras and speakers, UI: tablet (room navigator)	Robot + integrated tablet	PC/laptop + VR head mounted device	Laptop / Smartphone	Iphone/Mac + AR glasses
Cost	150.000 – 250.000 $	2600$	HMD: 400 – 2000€ Platform: free	Free	AR glasses: 600 – 1500 $ App: unknown
Type of interaction	GUI Touch (room navigator) and gesture (activated when someone enters the room)	GUI Touch (integrated tablet)	GUI and gesture (VR hand controls)	GUI	GUI / Touch and gesture

[Continues]

					
UX PROS	Human size image Photorealistic Very good image and audio quality Symmetric (local and remote user experience the same) Eye contact Hands free	Human height Photorealistic High resolution video and audio Hands free Physical presence and movable on the remote environment Easy to use by seniors	Immersive Shared (virtual) environment Human size image New and fun Symmetric Ability to interact together with virtual objects	Can see everybody in the same environment Photorealistic Increases engagement Can interact: high five, tap in shoulder Reduces background distractions	Human size image Photorealistic More immersive (no need of screen) Hands-free
UX CONS	Cannot interact with the remote context Restricted movement Requires complex installation	Requires space to store Asymmetric Small image Invasive (permitted user can activate it remotely. Local user does not need to "pick-up")	Avatar (cannot see real aspect, face expression or non-verbal language) Isolated from environment: Impossible to show real objects or activities No eye contact Not easy to use	Small people's representation Not very natural Cannot show or see details of real environment	Image quality still not very good No full body image AR glasses: No eye-contact, very difficult to have a symmetric experience When moving the image follows you floating (not natural)

2.3 Discussion

There is no doubt that video-communication is an expanding market and designers are working towards tools that satisfy the needs of specific use cases, as well as more immersive solution to make us feel closer to a face-to-face interaction. As previously mentioned, solutions oriented to work related environments or big events like conferences, receive more attention, but they are not very well adapted to the use at home for our personal communication and some exceed by far the budget of a regular user. One exception would be telepresence assistance robots, specifically designed for the monitoring of elderly at home or patients in healthcare institutions, but these systems are also limited to specific situations and budgets. Some of the common videoconferencing issues (like lack of shared context or poor support of nonverbal cues) might be addressed by 3D technologies, but the main drawback is their complexity for the general user, need of extra hardware and high cost.

From the results of this secondary research some factors should be explored further:

- Would users be willing to purchase a product which only purpose is to provide and improved personal video-calling experience?
- What is the attitude of users with average tech skills about the implementation of emerging technologies to everyday videoconferencing?
- How important is for users to see the real image of the person/s on the other side and perceive non-verbal cues?
- Can we provide with a more immersive and natural experience without the need of head mounted devices?
- How important is to be able to show details and activities from your own environment and see them on the other person's environment?
- What kind of features are more important when it comes to our private-life virtual meeting?
- What are the expectations of users on future video communication tools?
- Zoom fatigue during personal use of current tools

2.4 Research on user's needs

2.4.1 Goals

- Discover typical use patterns among users of videoconferencing tools for private/personal communication.
- Uncover needs and pain points they are experiencing with existing video communication tools.
- Know what kind of features users desire and expect from future social communication tools, and what is their attitude towards the application of new technologies on this area.

2.4.2 Methodology

Semi-structured interview: one-to-one interview where some of the questions are planned in advance, but the interviewer can also include non-planned questions depending on the course of the interview, in order to get valuable insights from the participant. This method was chosen to in order to get in depth qualitative information and its flexibility to ask follow up questions (Why and How). The initial script consists of 16 open-ended questions.

Online Questionnaire: This method was chosen to gather quantitative data from a representative sample of users. The online tool used was Google Forms and It consists of 24 items.

Participants

Table 2. Participant screener

Number of participants:	Interviews: 8	Online questionnaire: min.30
Age	18-65	
Gender	Balanced distribution	

Inclusion criteria	• To have a long-distance affective relationship (of any kind) • Frequent use of videoconferencing tools (at least once a week) to communicate with that person/s • Medium/average tech skills
Exclusion criteria	• Aversion to technology • Visual or auditory problems (after correction methods) • Use of videoconferencing tools only for work purposes

2.4.3 Procedure

Recruitment was done by posts on different social media (Facebook groups, Instagram and WhatsApp) where the inclusion criteria were explicitly mentioned. For interviews, each participant chooses a convenient timeslot from the ones provided. More information about the aim of the interview, what to expect and informed consent (annex) was sent by email. For the questionnaire, the link (Google forms) was already available in the post. Information and consent form were available on the first screen of the questionnaire itself.

Interviews (duration 20-30 min) took place online, using Microsoft Teams, a link to the meeting will be sent by email once the participant replies with the signed consent form. Most participants are people living abroad, that frequently use videoconferencing to communicate with family and friends. The interview starts with a brief introduction to establish rapport, followed by a series of open-ended questions arranged in a script (annex) and follow up questions if needed.

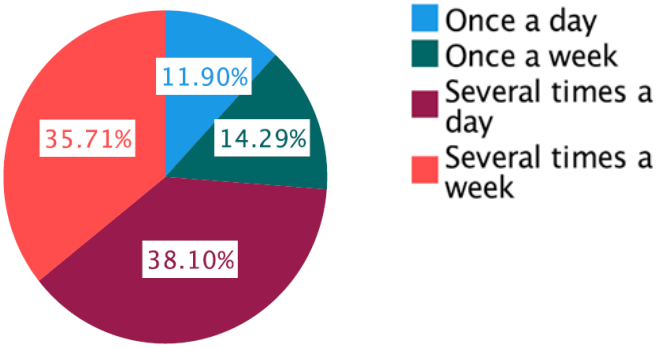
Questionnaire (duration 10 min) was made on Google forms (annex). It consists of 2 demographic items (age and gender), 4 Likert scale and 2 short-answer questions about their typical use of video-calling tools, 2 yes/no questions about their opinion on hypothetical new solutions, and 2 sets of items (Likert scale): first one about UX indicators (for mediated communication) and second one a standardized Zoom fatigue scale (ZEF) (Fauville et al., 2021).

Data analysis. Quantitative and qualitative analysis were performed. For the online questionnaire: Statistical analysis was performed using the Statistical Package for Social Sciences (IBM SPSS Inc.). Descriptive analysis was performed (central tendencies and frequencies). The ZEF score consists of the averaged ratings of the 8 items used, as indicated by the authors. To examine differences in ZEF scores per gender, a t-test was performed. Correlation among the different UX indicators was also examined. For the data obtained from the interviews, thematic analysis was performed, following an inductive methodology, by transcribing the content and using affinity diagraming technique.

2.4.4 Results

Questionnaire

42 respondents [25 female (59.5%), *M*age=34.57, *SD*age = 11.37].



Half of participants (50%) use videoconferencing systems for personal calls at least once a day or even several times a day.

Figure 6. Frequency of use of videoconferencing tools for personal meetings

Most of our participants (80.95%) made use of videoconferencing with family and friends more frequently than before during Covid epidemic.

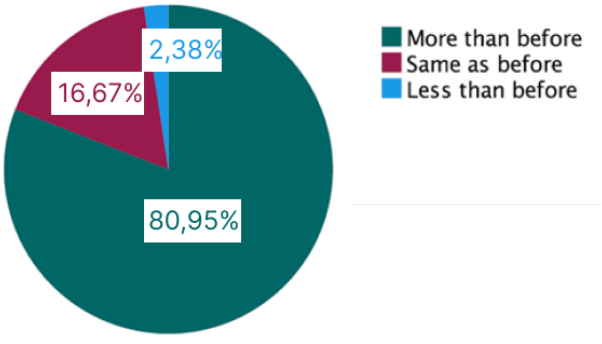


Figure 7. videoconferencing during covid crisis

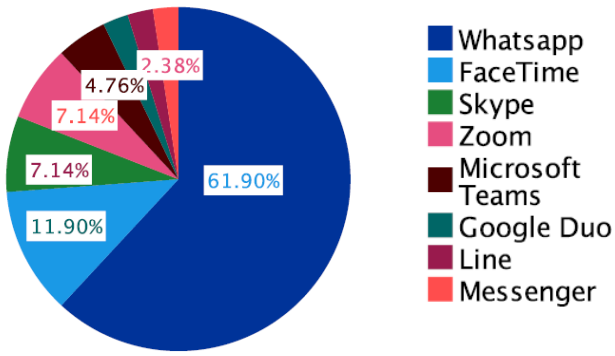


Figure 8. Tool of preference

61,9% prefer WhatsApp for their private videocalls, followed by far by FaceTime (11,9%). This means that, at least 61,9% of our sample use mainly their smartphone for this kind of calls since this feature is not available on the web version of this tool.

82% of our respondents Never or Rarely have personal group calls that involve more than 2 devices.

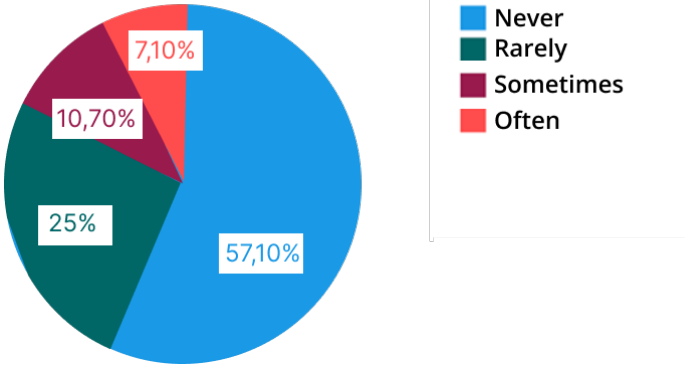


Figure 9. Frequency of personal group calls

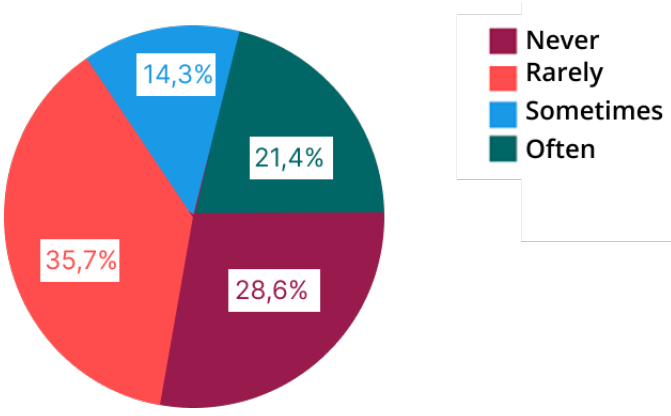
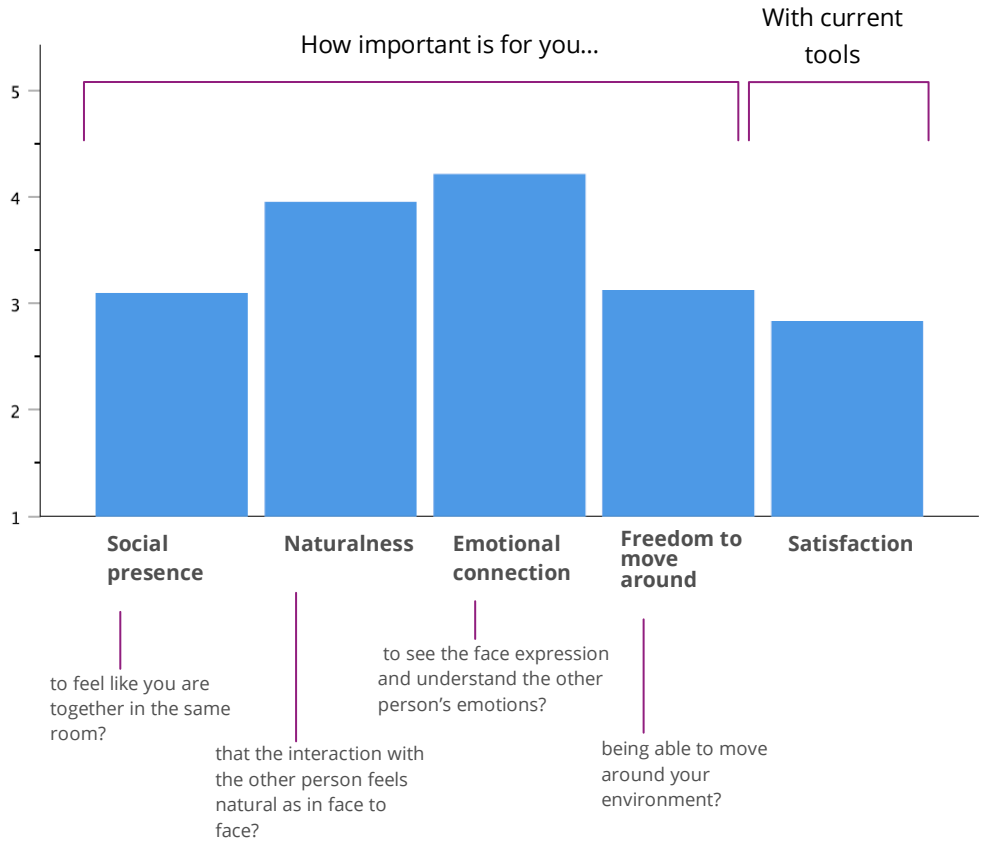


Figure 10. Frequency of videoconferencing for activities

50% of our respondents Never or Rarely use video conferencing to perform remote activities, while 21,4% do it often.

Most common activities are: sports at home, workshops/trainings, language lessons and coaching/therapy.



The UX indicator perceived as most important is the ability to stablish an *emotional connection* by effectively perceiving the other person/s face expression, followed by *naturalness* of the interaction. *Freedom to move* is perceived as moderately important. Users are **moderately satisfied** (2,79) with their current video-calling tools.

Figure 11. UX indicators for mediated communication + satisfaction

A further regression analysis confirmed **Naturalness as a predictor** of Social presence, Emotional connection and Satisfaction using videoconferencing systems.

Table 3. Mean ZEF score and mean scores per scale

Scales	Mean ± SD
ZEF Score	2.50 ± 0.84
General	2.63 ± 1.05
Social	2.42 ± 0.86
Motivational	2.60 ± 1.07
Emotional	2.33 ± 1.02

Average **Zoom Fatigue score** for our group was 2.50/5, which indicates a Moderate level of fatigue using videoconferencing in this context (personal).

However, it is worth mentioning that on average, the score of females (*M* = 2.76) was significatively higher [t (40) = 2.62, p = .01] than the score of males (*M* = 2.11) on this scale. No significative correlation was found between ZEF score and age.

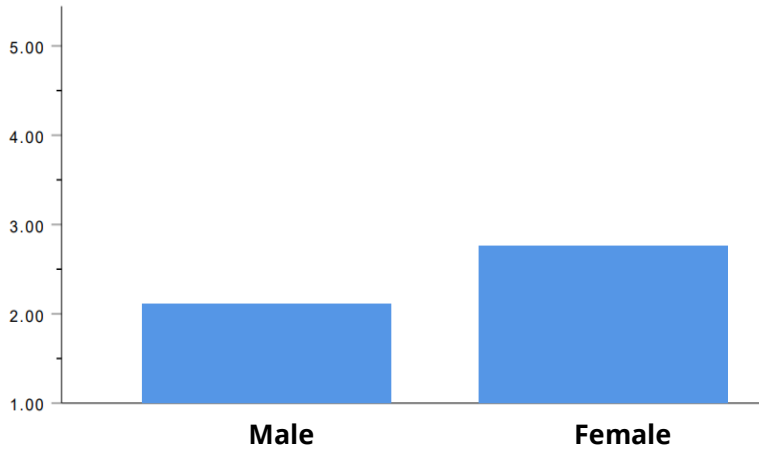


Figure 12. Average Zoom fatigue score per gender

After this finding, a chi-square test was performed, that indicated a significant association between frequency of use and gender: In our sample, men are using videoconferencing systems more frequently than women: **64.7% of men** use videoconferencing several times a day while only **20% of women** do it.

Two las questions were added to the questionnaire and answered by 28 respondents:

Would you purchase a product which sole purpose is to provide you with improved personal videocalls? (better quality / innovative features / more immersive)

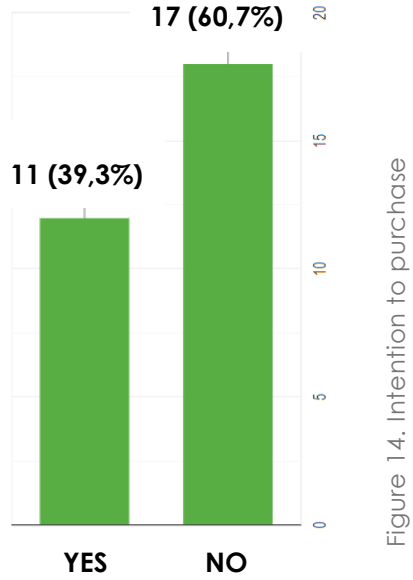


Figure 14. Intention to purchase

Would you dedicate a space at home for this product and purpose?

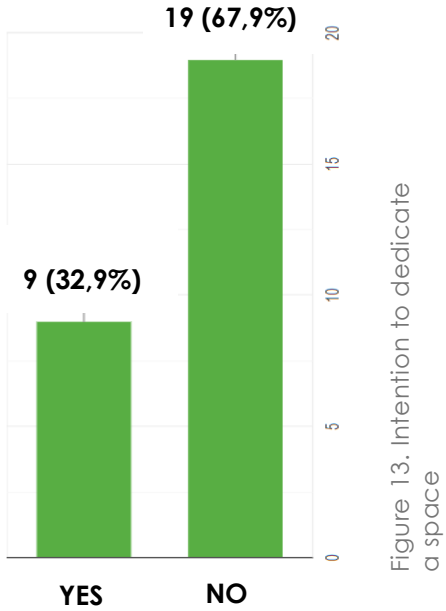


Figure 13. Intention to dedicate a space

Interviews

Table 4. Profiles of interview participants

	Gender	Age	Nationality	Occupation	Frequency personal VC	IT skills/attitude
1	Female	28	French	Designer	1-2 a week	Average, not very interested in new tech
2	Male	24	Indian	Engineer	Several times a day	High level, interested in new tech
3	Male	30	Spanish	Architect	2 a week	Average, “not afraid” of new tech
4	Female	35	Croatian	Engineer	1 a week (sometimes less)	High level, not very interested but confident on using new IT
5	Male	40	Dutch	Lawyer	everyday	Average, not interested or confident about new tech
6	Female	25	Italian	Translator	Several times a day	Average, “ok with trying new tech”
7	Female	34	Chinese	Graphic designer	3-4 times a week	Average, not very confident on using new tech
8	Female	54	Spanish	Sales manager	3-4 times a week	Average, not interested but accepts (with resignation) the use of new tech

Observations and insights gathered were synthetized and classified in categories by clustering them using an Affinity diagram.

Figure 15. Affinity diagram

Typical use of videoconferencing:

- Participants using videocall to communicate remotely with their romantic partner make use of videoconferencing more frequently than the others.

- Videoconferencing is becoming the regular way of communicating news and small details of daily life, over just-voice call or texting, but there are some situations that seem to be specially linked to communication with video like: talking about sensitive topics that imply an emotional response, important personal events like birthdays, seeing physical changes on the other person, showing an object or activity performed in our environment or vice versa.

> **"My sister had a baby, it is nice to see him growing".** P1
>
> **When I am cooking, I like to show what I am doing".** P2

- The most common device used is smartphone because of its flexibility and portability, mentioning the possibility to show their environment easily. Most common reasons to choose a VC tool are quality of connection, easiness to use, popularity among friends and relatives, and good usability on smartphone format (some mentioned that using Zoom or MS Teams on their smartphone was not nice). One decisive factor for some participants is the fact that WhatsApp is easy to use for their parents.

- Personal calls happen mainly at home, being important factors comfortability and privacy.

Evaluation of current VC tools

Main benefits (over other remote communication methods) identified:

- better interpretation of the message and anticipation of reactions, due to nonverbal language
- emotional connection

> **Video adds feelings to the conversation"** P5

- real-time connection
- possibility to "create international relations remotely" (P8) and reduce the perceived distance between people

Main pain points identified are:

- Impossibility to take part on the action on the other side and the other way around (not feeling present)
- Impossibility of physical contact and smell
- Lack of shared context
- Difficulty to keep focused on the conversation
- Emotions perceived as less intense
- Technical problems: lag image – audio, quality of connection
- No real eye contact
- Still requires planification

> **“If you are physically in the same place you have more opportunities to have small chats, remotely you have to concentrate everything in one call”.** P8

Also, two participants mentioned fatigue: “I have to use it a lot for work, so using it again for personal calls makes me feel tired… looking at a screen for a long time is tiring” (P3).

When asked about the influence on their long-distance relationships, they all agree on videoconferencing being a useful resource to maintain and strengthen their relationships while being apart and some of them mentioned how useful it was during Covid period

> **When you cannot see them, distance feels bigger”** P7

but most consider this is the case just when meeting in person is not possible and VC will never replace face to face interaction.

Expectations about future videoconferencing systems

Futuristic movies are a common reference to describe future videoconferencing systems, and most of our participants included the word “hologram” in their description.

From several comments it can be extracted that screen-less systems are a common expectation. Also, frequent concepts were: feel, touch, smell and presence. All assume that these systems will inevitably evolve and lead to different ways of communication in the future, making it more socially normalized. Some interesting ideas were introduced: “being able to walk with the

person", "doing things together, like having a coffee break", "being able to see the person in your own environment or being immersed in the other's environment depending on the situation".

Expected characteristics for VC systems based on current mental models (popular apps available):

- Intuitive interface
- Not feeling all the technology and "electronic stuff" around
- Quality of image and sound
- Responsiveness on different devices
- Easy functionality check and troubleshooting
- Lite (not very data-consuming so it can be used without Wi-Fi)
- Free of charge
- No need of email registration, links or scheduling upfront (allows flexibility and improvisation)
- Allows fluid group calls

> **"MS teams or Skype are for professional calls, because you need to register with an email address, schedule a meeting… for personal calls it needs to be easy, you need to be able call whenever you want"** P8

Attitude towards application of new technologies to personal communication

Intention to use a system that "allows you to feel as if you were in the same room with the other person" was high (8.6 /10). Some considerations expressed: "It can potentially erase the need to travel" (P1), "It would be valuable for long distance (romantic) relationships" (P3).

As with any new technology, recommendation from a person of trust is an important factor to embrace it, as well as the experience of people on the other side of the call,

> **"If they are not easy enough for older people like my parents that would be a problem".** P4

therefore, inclusive design should be implemented.

Discussing VR/AR more specifically, most participants identified potential benefits ("They would help to feel the person closer", P2) and recognized that these technologies align with their idea

of future video communication. On the other side, there is a generalized perception of XR as "one time thing" associated with entertainment/gaming situations and a conservative attitude when it comes to its use for remote communication at home (especially about VR): "I would not see myself putting a headset on for a normal personal conversation, it is too complicated" (P1). Some comments were made related to human representation in VR/AR, agreeing on the need of a real image of the other person rather than using an avatar for the communication to be satisfactory.

Lastly, the most repeated concern expressed on the use of this kind of technologies in this context is data privacy ("what information will be collected from your environment. How much they know about you", P1), followed by psychological effects.

2.4.5 Insights

It is obvious that videoconferencing (VC) plays an important role in maintaining remote relationships and it is taking over other communication methods by allowing users to **share emotions, non-verbal details and sensitive moments** with their loved ones. This has been evidenced by an increase in its use during covid pandemic. Because of the intimate and informal nature of these interactions, VC tools designed for the workplace could not be just translated to this context. Our solution must be flexible and adapted to the most common activities that users perform and characteristics of private-life videocalls: one-to-one or small group calls, sports, language lessons, workshops, sensitive conversations where face expressions are key, etc.

The fact that the most typical app for personal communication (WhatsApp) is chosen largely because of its easiness to use across generations, leads to think that the UI must **be intuitive and accessible regardless of age or technological skills**. Being smartphone the most typically used device, we could think that portability is also an important characteristic, but in fact, most of these personal calls are performed **at home because of privacy and comfortability.**

Feeling of naturalness, defined as the extent to which users perceive the interaction as predictable, logical or in line with their expectation (Skalski et al. 2011) of real-life communication, seems to be an important factor to achieve a higher sense of social presence and emotional connection. Ways to enhance the feeling of naturalness, such as real eye contact, effective turn taking, synchronism (image-audio), natural/organic UI or life-size image of the other person must be explored.

Expectations towards future systems have to be interpreted carefully, since mental models play an important role: they are quite limited to the possibilities of popular video communication apps and it's difficult to envision other features without falling into unrealistic scenes from sci-fi movies.

Some of the **pain points** identified are relative to the lack of stimulation of other senses like smell or touch, or the impossibility to interact with the environment of the other person/s or do physical activities together. The first has been quite explored (Huisman et al, 2013) but there are no commercialized solutions and it is a whole other area of research that will not be included in this project. The second area, has been mainly explored in the context of collaborative VR (Auda et al. 2021) or remote robot handling. Telepresence robots or systems that allow the user to have some kind of physical presence and agency on the other side are an interesting option to take into account.

Zoom fatigue is a phenomenon that should be considered when developing any new VC system. The fatigue caused by an increased number of remote work-meetings affects user's attitude towards using screens for their personal communication. The fact that women experience more Zoom fatigue than men, and videocall less frequently, can be partly explained by mirror-anxiety (Fauville et al., 2021). This can be triggered by the self-view in video conferences, related to a higher awareness of being observed and higher self-consciousness in women than men during videoconferences. Therefore, removal of self-view could help in reducing fatigue.

Implementation of **new technologies like VR or AR** in our daily life still sound futuristic and unreal, but people know them and see potential benefits. One fundamental part of remote interaction with people we trust is to be able to share part of our physical environment (and vice versa), for this purpose AR seems more appropriate. Also, VR comes with the complexity of using a heavy headset and doesn't provide the "naturalness" that users are looking for this kind of communication.

The answers on **intention to purchase**, lead to conclude that, while the general user has interest on improving the quality and experience of video calling, is not willing to purchase a product which only purpose would be that one, less to dedicate a space at home specifically for it. Therefore, in my opinion, this solution should be based on popular devices that can be used for other purposes (smartphone / tablet / smartTV) with the possibility of adding (small) accessories in order to improve the experience.

Finally, if we want a communication system for domestic use to be successful, users must be assured **privacy and effective data protection**.

Definition

Insights gathered on desk research and research with users have been synthetized using personas and user journeys. First technique gathers demographics and potential goals, motivations and frustrations of typical user profiles, whereas the second represents a potential use case related to these personas, where contextual details, actions, emotions, thoughts, pains and gains are covered. Both techniques allow a visual representation of insights that facilitates extracting aspects that are key for a satisfactory user experience and areas of improvement, in order to define requisites for our prototype and opportunities to explore for the future design, as well as sharing this information with different stakeholders.

3.1 User personas

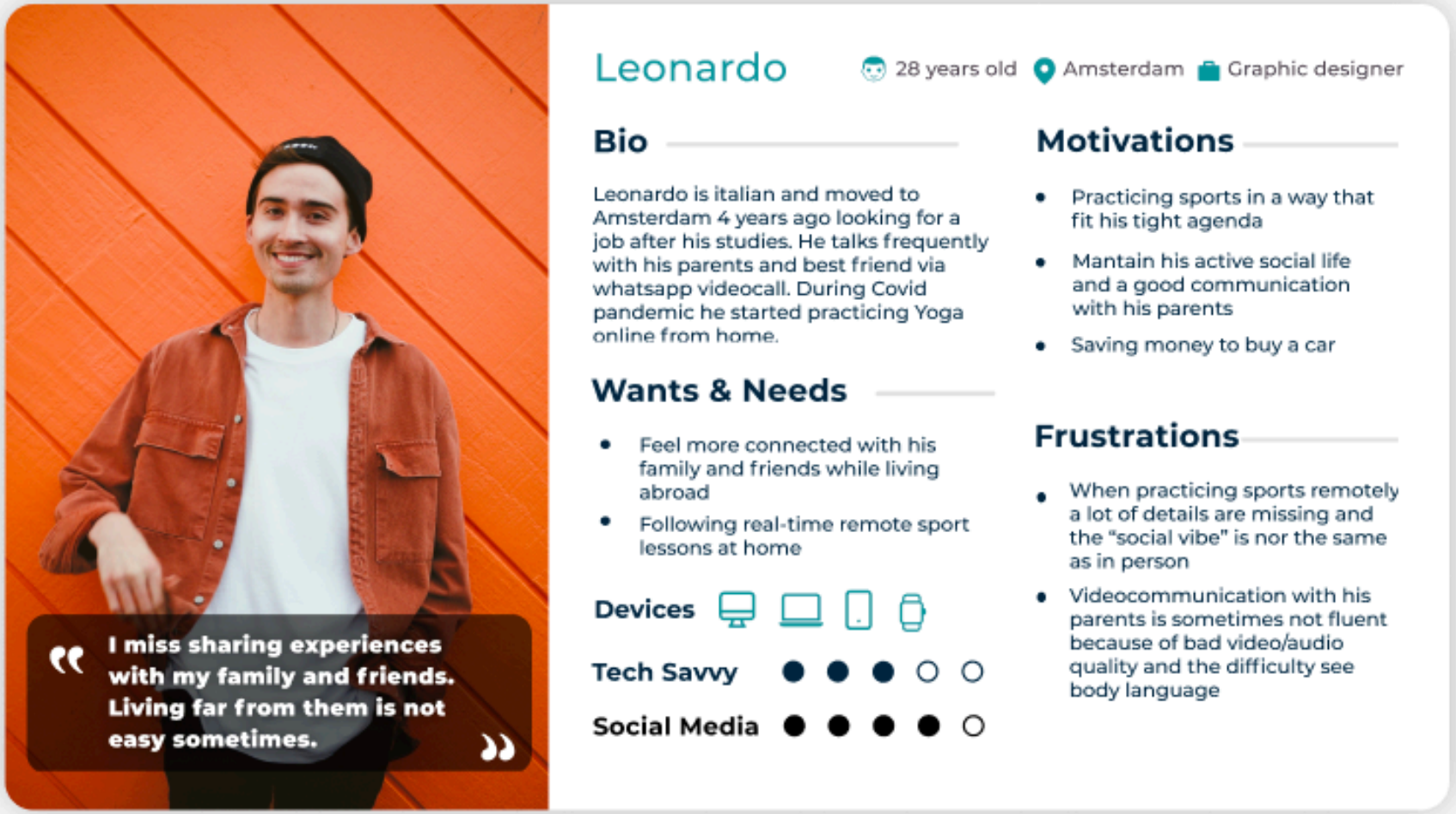


Figure 16. User persona I

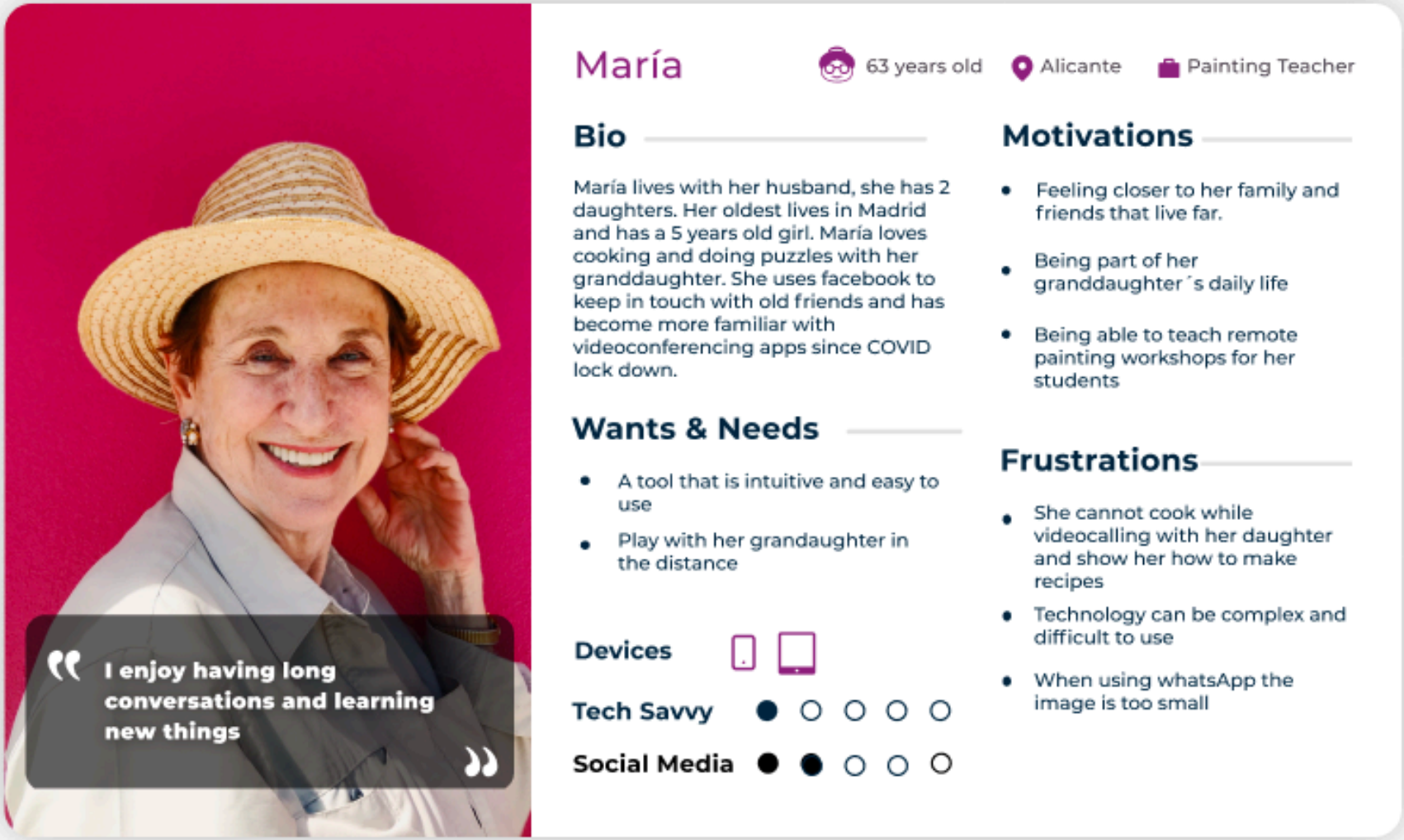


Figure 17. User persona II

3.2 User Journey maps

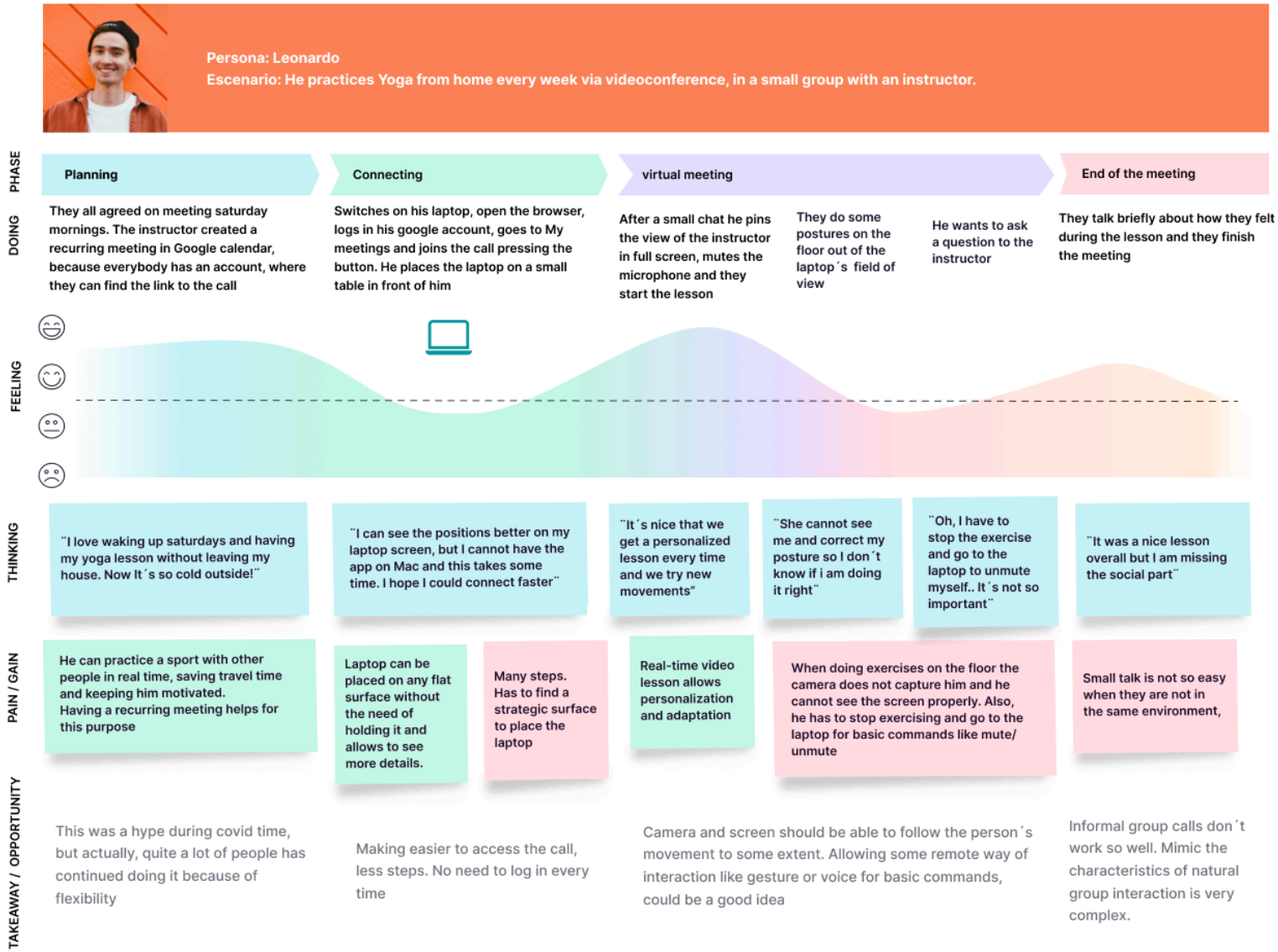


Figure 18. User journey map: Leonardo

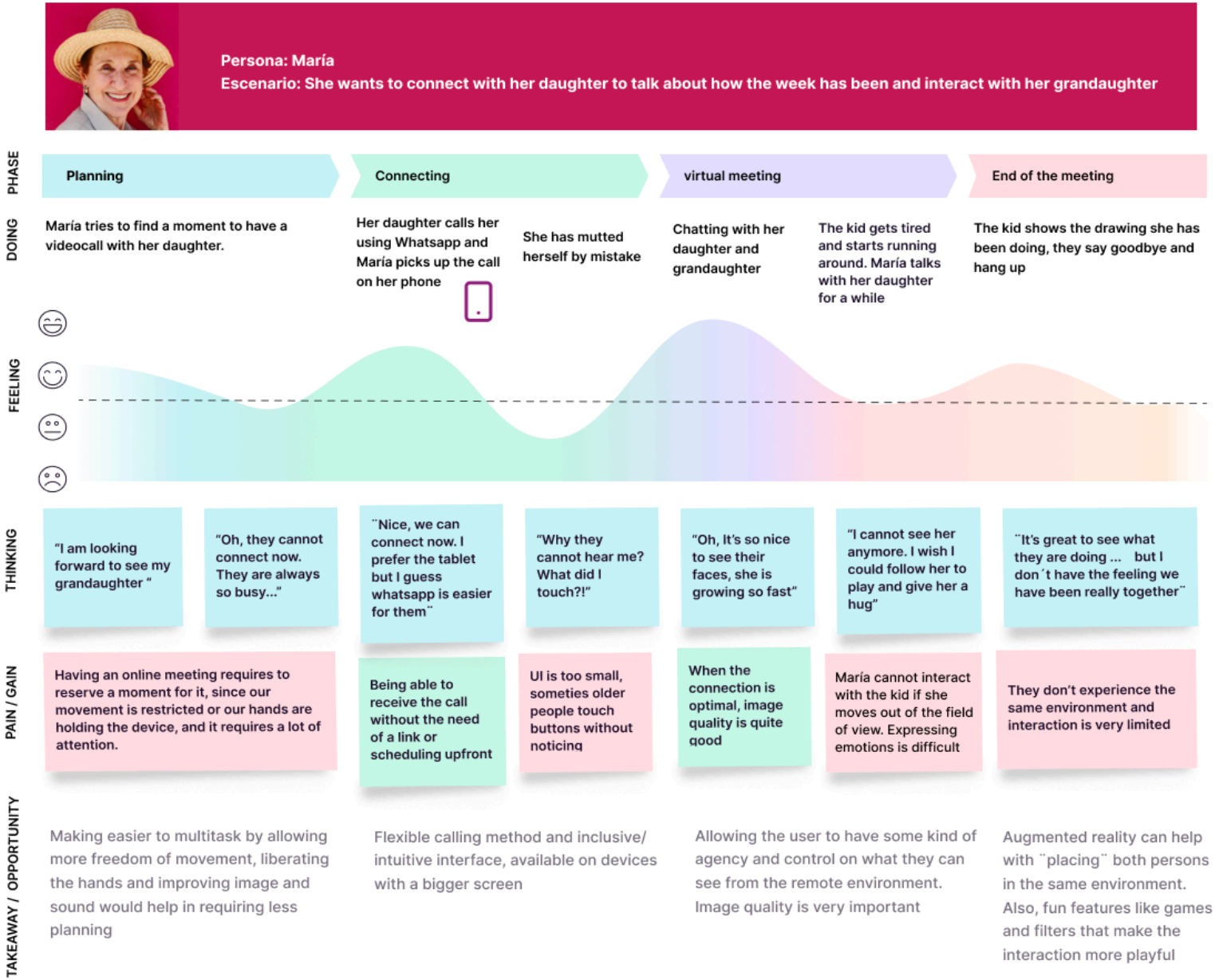


Figure 19. User journey map: María

3.3 Design requirements

FUNCTIONAL

- Flexible any-time calling (no need of link or invitation)
- Camera and screen follow the person´s movement to some extent to allow being captured and screen visibility when moving
- Possibility of zooming image in and out
- Voice/gesture control of basic commands
- Fun features like filters and games
- Best experience on one-to-one calls but allows small groups.
- Allows to record calls and send video recorded messages
- Different calls, different needs: Allows to choose between viewing the real environment on the other side or providing the illusion of being immersed in the same environment (AR)*

INFORMATIONAL

- Easy functionality check and troubleshooting
- Clear instructions are shown on the screen when an action is required from the user for a better performance

NOT FUNCTIONAL

- Provides a more natural social interaction through some kind of embodiment on the remote environment
- HD image
- HD audio
- Intuitive interface with good contrast and big enough typography and buttons, to make it accessible for older people.
- Easy installation and connection
- Effective encryption and data protection
- Removable self-view

ENVIRONMENTAL

- It doesn´t require a lot of physical space (based on existing devices + possible small accessories)
- Doesn´t require to buy expensive hardware
- Free app

*The feasibility and effect of this feature will be explored trough a proof-of-concept experiment, with a prototype developed by the independent research organization TNO, and will be explained in the next section

Exploration

4.1 Proof of concept

As previously mentioned, the application of emerging technologies, such as extended reality, to video-communication (VC) systems is being explored and might be effective in providing an improved user experience and bringing people closer. In this line the Dutch independent research organization TNO has developed a prototype of a VC tool based on photorealistic augmented reality (AR), from previous research on social XR (Gunkel et al., 2019). To explore the effect of this technology on the user experience, in comparison with common VC tools, we have performed an experiment that will be the subject of this section (complete study in the annex).

4.1.1 Introduction and procedure

For a successful application of new technologies to personal communication tools, especially for those being used in a more intimate and emotional context (domestic), it is fundamental to put the users in the center. Though it is difficult to get a universal definition of UX for social immersive technologies, specific factors involved are *naturalness of the interaction, social presence* and *interpersonal closeness* (Li et al., 2019).

> **“ The assessment of satisfaction with entertainment systems in teleconferencing and collaborative virtual environments is based largely on the quality of the social presence they afford.**
>
> - (Biocca et al., 2001)

Social presence is defined as the sense of "being with another" or, in the context of mediated communication, "the perceptual illusion of non-mediation" (Biocca, 2003), and seems to have an impact on long-term emotional connection (Gooch, 2015). Therefore, a satisfactory telepresence system would have the potential of supporting and strengthen long-term personal relationships which can add an enormous value in times of crisis or other circumstances of separation from the loved ones.

Goals:

- To compare TNO AR-based system with a regular videocall (MS Teams) in terms of perceived **social presence**, **closeness** and **global UX**.

- To explore the attitude towards the AR solution and other factors that can improve the user experience of mediated communication systems.

Logistics:

The experiment took place in the lab, using two contiguous rooms. Recruitment of participants among TNO employees, their acquaintances and students. Rooms were arranged to look cozier so people feel more comfortable having a personal conversation.

Participants:

36 participants (50% female), 9 pairs (friends, couples, family, close colleagues) per condition. Each pair of participants was randomly assigned to one condition only (between subjects)

Method:

Each participant in the pair is located in one room, they are instructed to perform a conversational task (plan a trip together) using MS Teams (condition I) or the AR tool (condition II) their behavior is being recorded to study behavioral patterns. After the conversation, every participant answered a questionnaire, about their experience. The duration of the experiment is ~30 min.

Evaluation:

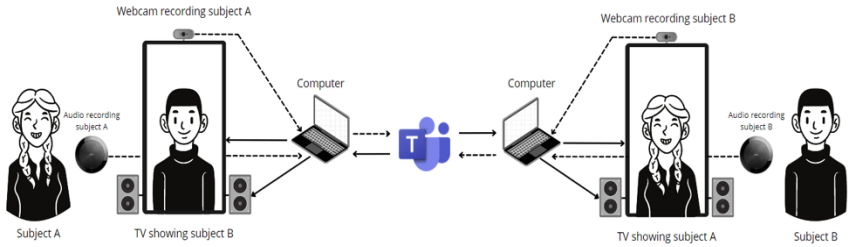
For this study we designed an evaluation methodology based on multiple existing questionnaires, objective measure (behavior) and some questions included by ourselves:

Social Presence	**Interpersonal closeness**	**Global UX**
• H-MSC-Q of which only the social presence scale will be used (Toet at al. 2021) • Networked Minds Questionnaire (Biocca et al., 2001)	Other in Self (IOS) scale (Aron et al., 1992)	The User Experience Questionnaire (UEQ) adapted
Behavioral patterns	**Attitude towards AR for VC**	
For this purpose, conversations were recorded (participants gave their consent)	Intention to use (score) and open questions	

4.1.2 Technical setup

Both setups are symmetrical for both participants. We chose a large screen (46 inches) to provide a human-size image. Many previous studies have suggested that a life-sized view is likely to enhance the user's sense of social presence during a videoconference (Ahn et al., 2014; Ishida et al. 2011; Koh, 2010). Same display was used in both conditions to isolate the effect of the AR-based image.

Condition I. Regular videocall (MS Teams)



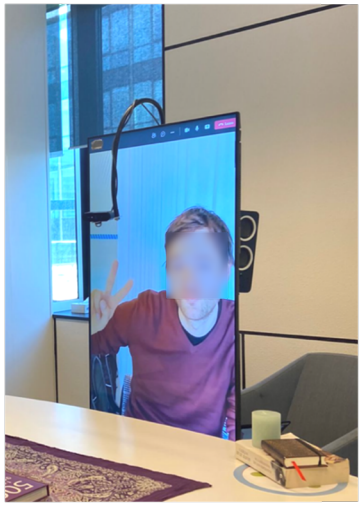


Figure 21. Videocall using MS Teams

Both participants see a 2D image of the other participant and his/her background. Self-view will be disabled to avoid distractions and maintain both conditions similar (the AR condition does not have this feature).

Condition II. AR-based videocall

Figure 20. AR-based videocall

The participant was recorded using a Kinect camera. Attached to the back of the screen is a webcam which provides the background image. The videocall takes place using the AR based tool. When looking at the screen, participant A sees participant B “projected” (3D point cloud image) onto the background image of the room participant A is in, so that he/she appears to be sitting opposite in the same room.

4.1.3 Results

Analysis

Statistical analysis was performed using IBM SPSS Statistics 28 (www.ibm.com). A significance level of 0.05 was used for all hypothesis testing. Thematic analysis was used to analyze qualitative data from open questions.

Questionnaires

Social presence:

NMQ (Networked minds. Biocca et al., 2001) & HMSCQ (Holistic Mediated Social Communication Questionnaire. Toet et al., 2021) did not show differences between conditions. Both conditions received a rather high score (above 4/7 on the NMQ and above 5/7 on the HMSCQ).

Table 5. HMSCQ scores

	MS Teams communication	AR communication	*p* value
Total score HMSCQ	5.60 ± .81	5.50 ± .65	.60
Copresence scale	5.50 ± 1.14	5.00 ± .95	.76
Intimacy scale	5.50 ± 1.13	5.50 ± .99	.56
Credibility scale	5.50 ± .74	5.00 ± 1.32	.01
Reasoning scale	6.00 ± .69	6.00 ± .79	.79
Behavioral scale	5.75 ± 1.05	6.00 ± 1.13	.93

Closeness

No significant differences were found between conditions.

Analysis of correlation between social presence and interpersonal closeness (IOS) revealed a strong positive correlation (r = .72, p = <.01) when social presence is measured with the new HMSCQ.

Table 6. UEQ Benchmark comparison

	Scale	Mean	Comparisson to benchmark
Video-call system (MS Teams)	Attractiveness	1.83	Good
	Stimulation	1.55	Good
	Novelty	0.51	Below Average
	Dependability	1.17	Above Average
AR system	Attractiveness	1.83	Good
	Stimulation	1.25	Above Average
	Novelty	0.82	Above Average
	Dependability	0.97	Below Average

Global UX

No significant differences were found among systems. Nevertheless, it is worth mentioning that, according to the authors of this test, scores > 0.8 represent a positive evaluation. Therefore, both setups have received a quite positive score, also using the benchmark provided by the authors.

Behavior

Behavior was coded into 7 factors:

1. Non-verbal communication: Hand gestures
2. Natural postures and body movements (as in face-to-face): such as leaning head on the hands, crossing arms, etc.
3. Interaction with objects
4. Showing objects to the remote person
5. Laugh
6. Leaning back relaxed (duration)
7. Engagement: was the conversation interrupted at the end of the experiment?

There was significant difference in non-verbal communication, t (31) = 2.3, p = .03. **Participants used more hand gestures to communicate using the AR tool** (M = 11.88, SD = 5.63) than using MS Teams (M = 6.94, SD = 6.48). There was also a significant difference on the number of times they laughed and the time they spent leaning back relaxed, higher using the AR tool.

It was observed that the interaction in both conditions looked, in general, more **natural and dynamic** (as in face to face) than what we are used to see when using popular videoconferencing apps on small size devices (smartphone/tablet/laptop).

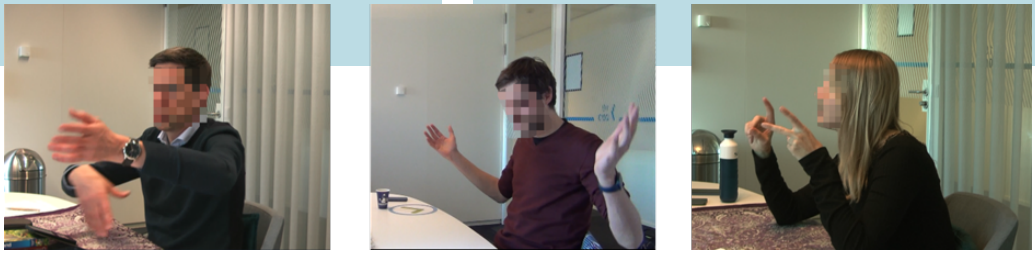


Figure 22. Participants using hand gestures (non-verbal communication)

Attitude towards the AR-tool: qualitative analysis

ITU: Participants assigned to the AR-based communication system, were asked to what extent, on a scale from 1 (strongly disagree) to 7 (strongly agree), they would use this technology again for their personal calls. Mean score was **5.44** which indicates a rather high intention to use.

This score was supplemented with comments about pains, gains and suggestions about the tool, gathered with three open questions:

What did you like about the system?
What would you change from the system?
Considering we are designing a system for domestic-personal communication, do you have any suggestions?

Gains

- Human-sized image
- More Natural
- Less tiring
- Feeling connected
- Closer to Face-to-face
- Same room (background)

Pains

- No real eye-contact
- Image distortions
- Resolution
- Audio lag
- Background-Person integration (edges)
- Inability to move (location in the room)
- Poor image quality when moving

Comments/suggestions for improvement

Image quality and size:
"The image edge fragmentation was a bit distracting…" P22

"I wonder whether the main point is just the "size" of the image. That's what makes the biggest impact for me." P24

One-to-one meetings:
"I'm not sure how this would work in a group setting, but for one-on-one calls I think this is a big improvement for communication as a deeper connection can be made." P26

Better than current tools for a relaxed conversation:
"I would not prefer this above a physical meeting, but if it is more practical to have an online meeting, this is a great opportunity for a relaxed atmosphere to talk to each other. Better than a Teams or Zoom meeting." P33

No need of 3D image:
"The quality of the person's edges. The 3D capture is unnecessary here; a 2D capture with background removal would suffice." P32

4.1.4 Conclusion

W**e cannot conclude that the new AR communication tool provides with more social presence, but it appears to induce a more natural interaction (as in face to face).**

Both scenarios actually scored very positively on social presence and general UX, as well as, promoting a more dynamic and natural-looking conversation than what we can observe during common VC on small size devices. This leads to consider that **other factors** present in our experiment are having an influence. Such as:

- being hands-free, allowing them to gesticulate more
- not having a self-view, does not restrict their hands/arms movements (so much) to the camera's capture range
- having a human-sized image, with good-enough (image and audio) quality, provides with more realism and allows the person to lean back comfortably, while keeping the "eye-contact", not missing details

Therefore, while the implementation of **extended reality** has potential to improve the naturalness of this interactions, to achieve a more advantageous holistic UX for video-communication, it must be **integrated in a good strategy involving other functional, technical and contextual factors.**

4.2 Redefinition & Conceptualization

This proof of concept has provided with new insights that supplement the ones obtained on the exploratory research on user's needs. This information helps to improve the definition of requirements, as well as, the conceptualization of the system and strategy to follow. An overview can be seen on the following concept map of the system:

- Concept map

FUNCTIONAL

- Flexible any-time calling
- Fluid experience between devices for different scenarios/goals
- More natural interaction that allows mobility: Face tracking camera / smart motion tracking (AI)/ screen visibility when moving
- Possibility of zooming remote image
- Fun features like filters
- Best experience on one-to-one calls but allows small groups.
- Allows to record calls, send video recorded messages and text messages

INFORMATIONAL

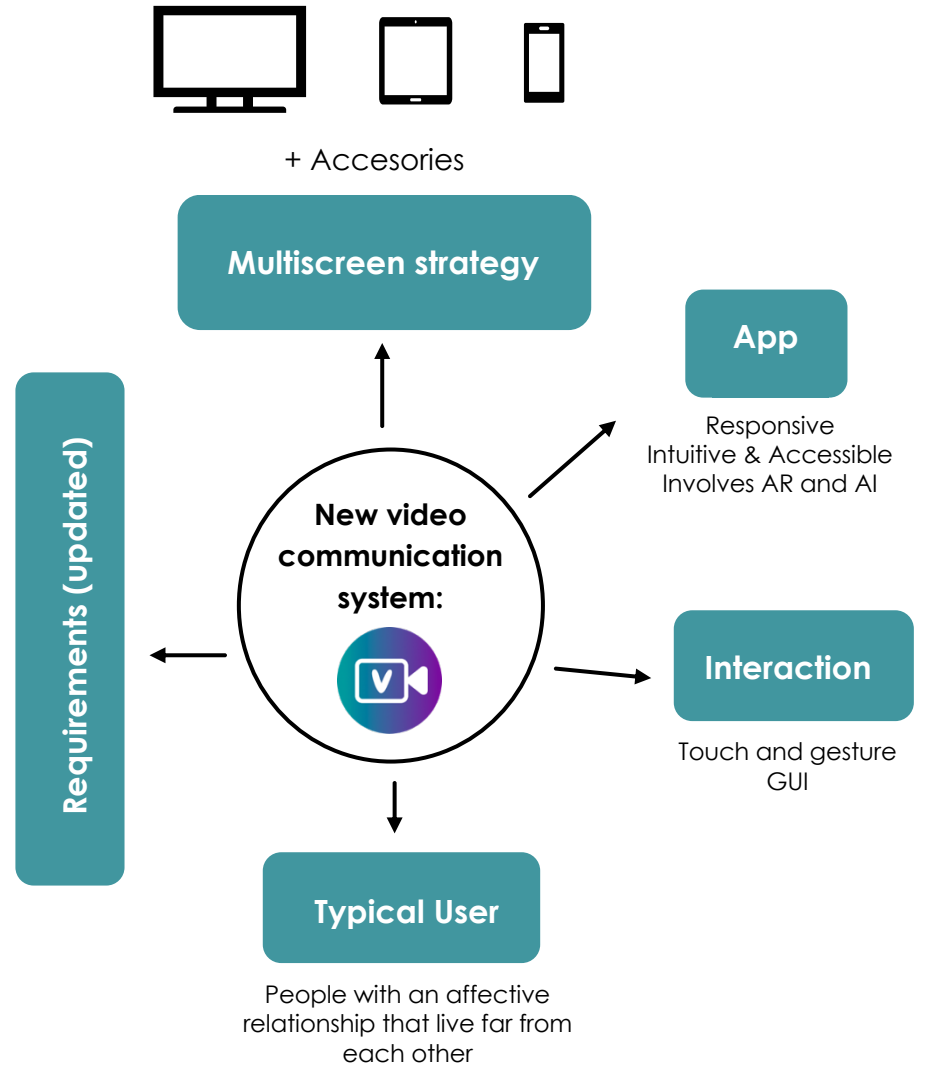
- Easy functionality check and troubleshooting
- Guidance when using it first time
- Clear instructions when actions are required and information about the status of the system
- Inform when connection issues
- Sight off controls (appear when touching the screen)

NOT FUNCTIONAL

- Increased sense of social presence: AR feature to provide the illusion of being immersed in the same environment (option when using a portable display). Human-size image possible
- HD image
- HD (directional) audio
- Good synchronicity image-voice
- Intuitive interface accessible for older people.
- Easy installation and connection
- Automatic dropping video quality to preserve audio quality when poor connection
- Automatic reconnection
- Effective encryption and data protection
- Removable self-view
- Calls without time limit

ENVIRONMENTAL

- Flexible-adaptable to different environments
- It doesn´t require a lot of physical space
- Doesn´t require to buy expensive hardware (based on existing devices)
- Free app



LIMITATIONS

- Impossibility to combine AR (person projected onto real background) with a human size image on a domestic scale. Requires expensive new hardware of big dimensions.
- Real eye contact very difficult to achieve. Limited to existent hardware: location of the camera is always above the eyes.
- Image quality is dependent on connection quality.

Creation

This section includes the design and visualization of the system, based on previously defined requirements and conceptualization. It contains the information architecture, way of interacting with the app and system as a whole, as well as the appearance of the user interface, using sketches and a functional prototype that will be tested in a later evaluation phase.

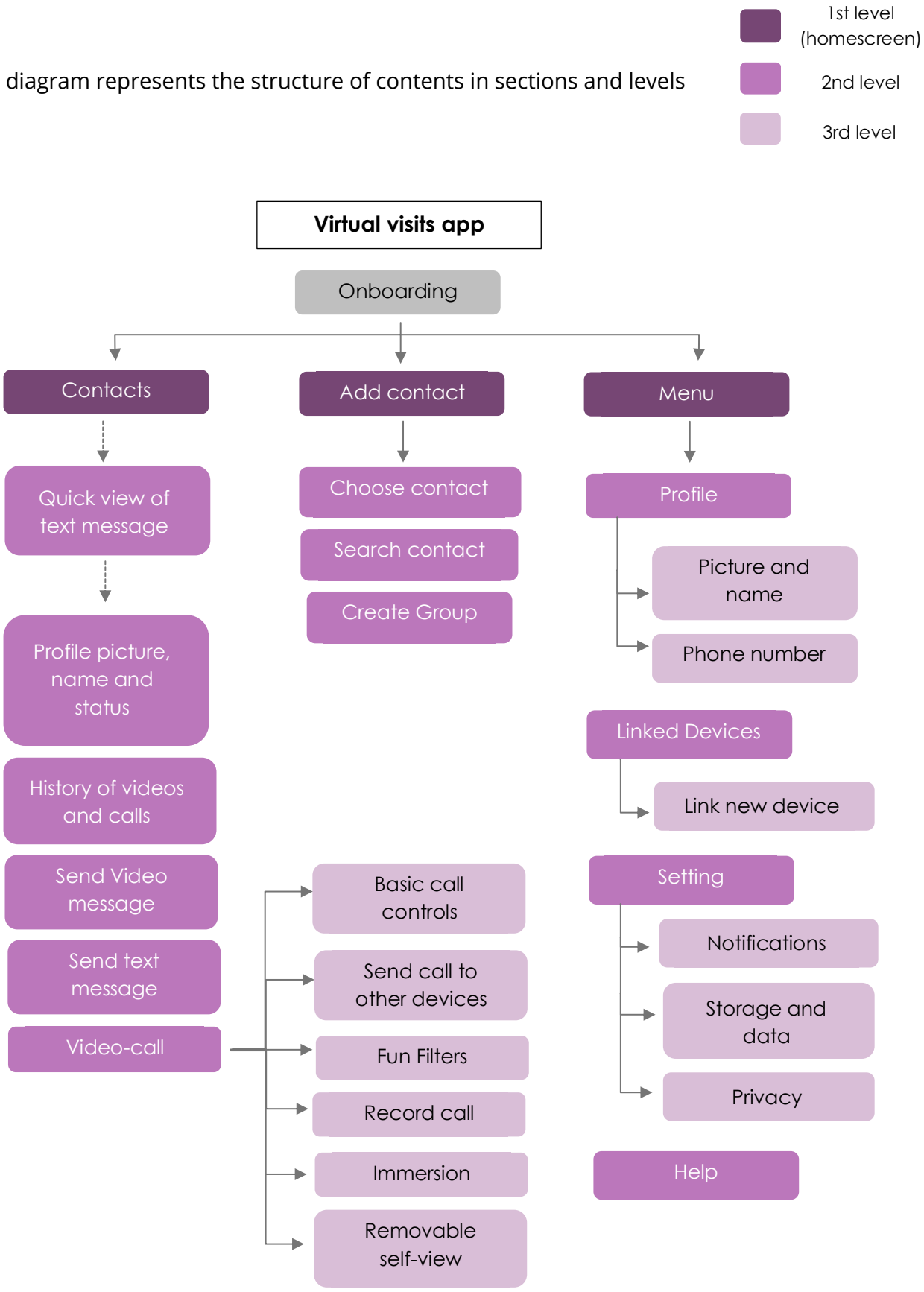
5.1 App content inventory

Following contents and architecture correspond to features that will cover some of the requirements stablished, as well as standard features for this type of apps, and it's been inspired by the most popular app for our typical users: WhatsApp, to match their existent mental model.

- **Onboarding**
- **Fast access to contacts from home-screen**
- **Add contacts by phone number** (search in phone agenda)
- **Create group**
- **Quick view of text messages received**
- **Contact information**
- **History:** Text messages, recorded calls, sent and received videos
- **Call controls:** answer, finish call, switch off camera/microphone
- **In-call options:** send call to other devices, fun filters, record call, immersion feature (AR)
- **Menu:**
  - **Profile:** Name, picture, phone number (change phone number)
  - **Linked devices:** existing devices linked, link new device
  - **Settings:** Profile, Notifications, Storage and data, Privacy
  - **Help**

5.1.1 App content tree

This diagram represents the structure of contents in sections and levels



5.2 Design of interaction

For a better understanding of the interaction with the system this section starts with a sketched visual representation of its components:

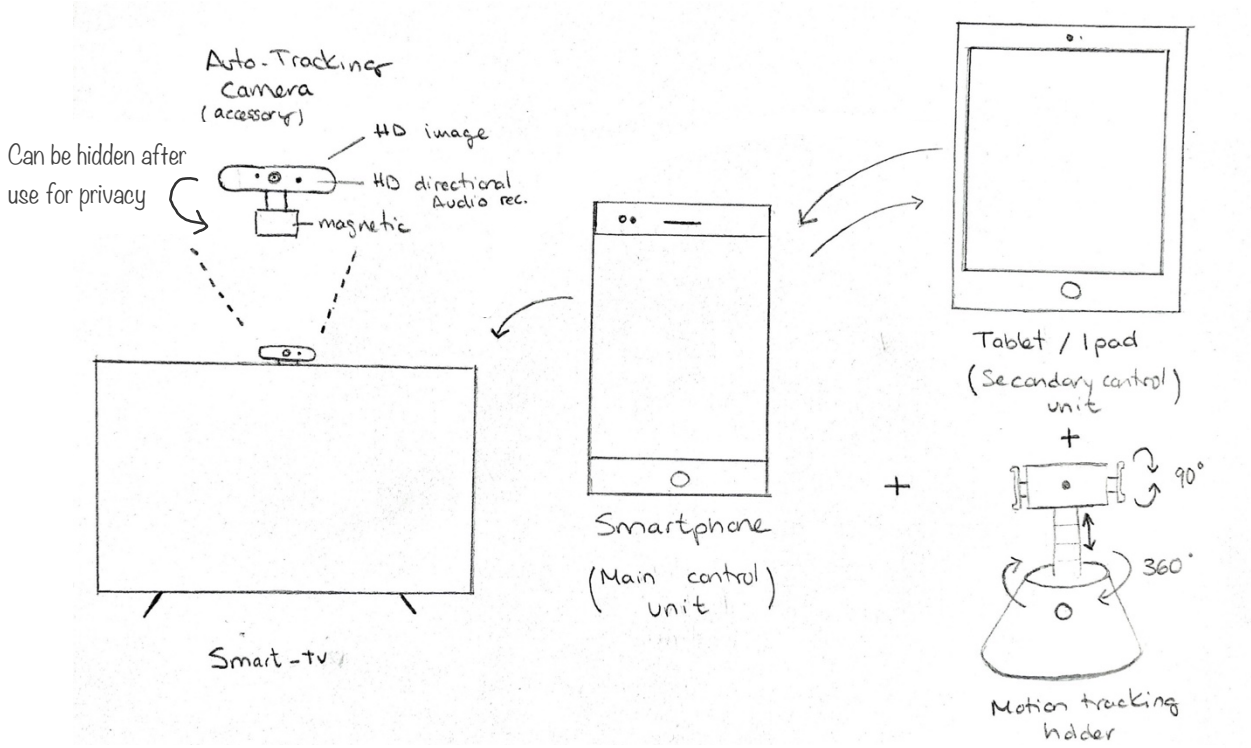


Figure 23. Sketch of system elements

The style of interaction that correspond to our system is *touch* and *gesture*:

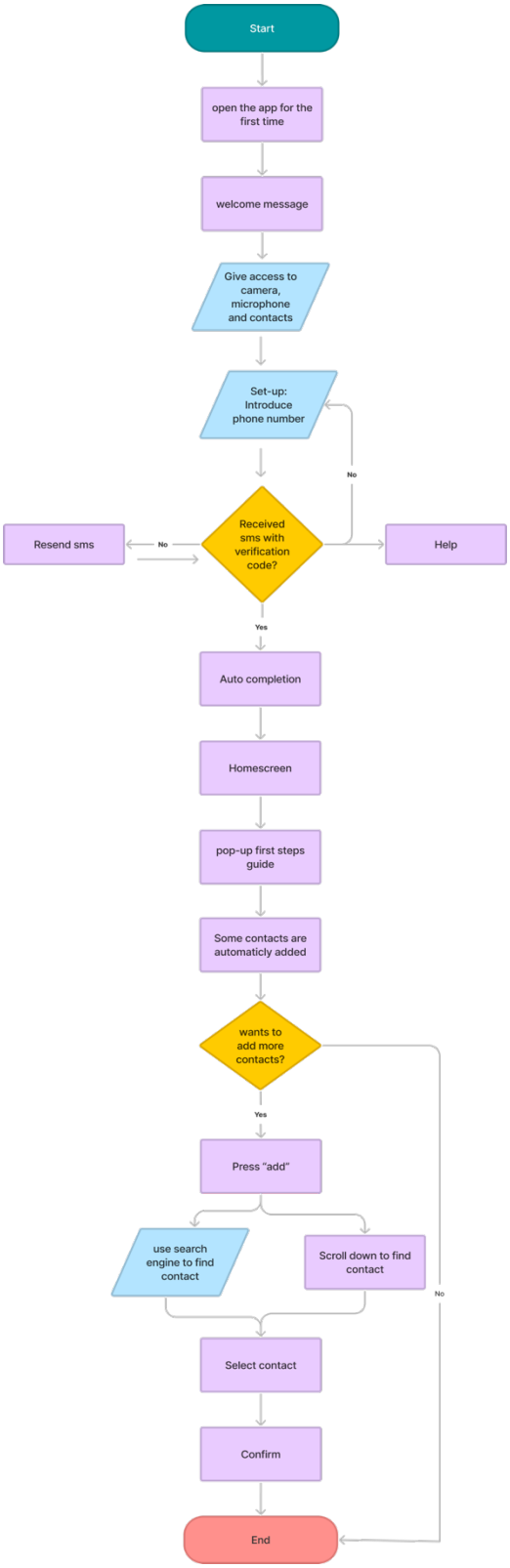
- **Screen-touch:** for input devices (GUI): Smartphone (main) and tablet (secondary)
- **Manipulation:** place portable device and switch-on tracking holder.
- **Natural interaction (gesture):** despite natural gestures and movements while having an online meeting don't a have a specific goal towards our system, motion tracking devices capture them and send feedback (follow their movement), therefore a natural interaction occurs.

- To have an effective cross-device experience, all devices must be connected to the same network with the app installed.
- When transferring a call to a smart TV, call controls will remain on the smartphone to make it easier to interact at a distance. When transferring to tablet, the call can be controlled by both smartphone or tablet.
- Immersive feature (AR) can be used for videocalls or sending video messages.
- Voice and gesture-based (touchless) interaction are not being considered for now since they could interfere with the normal course of the meeting. Designing this type of interaction would require further research.

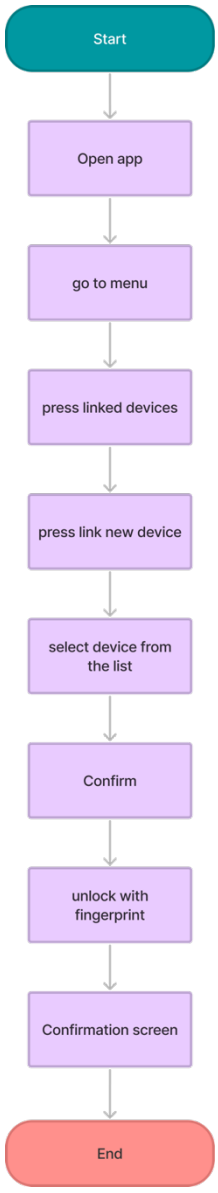
5.2.1 Flow charts

Based on the previous tree of contents and the conceptualization of the system, this flow charts will help to get an overview of the user's navigation with our product:

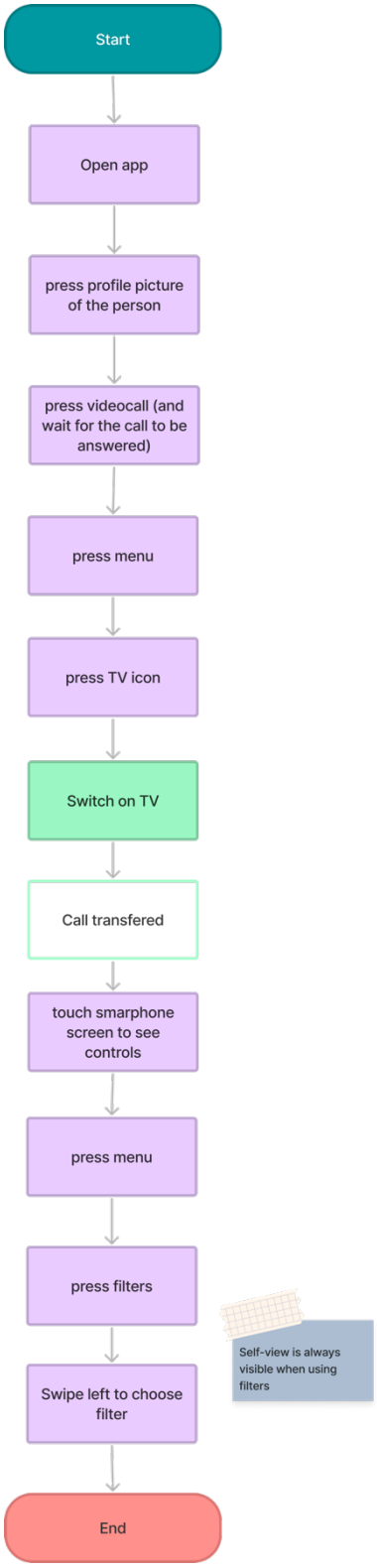
1. Onboarding. Install, configure the app and add first contacts.



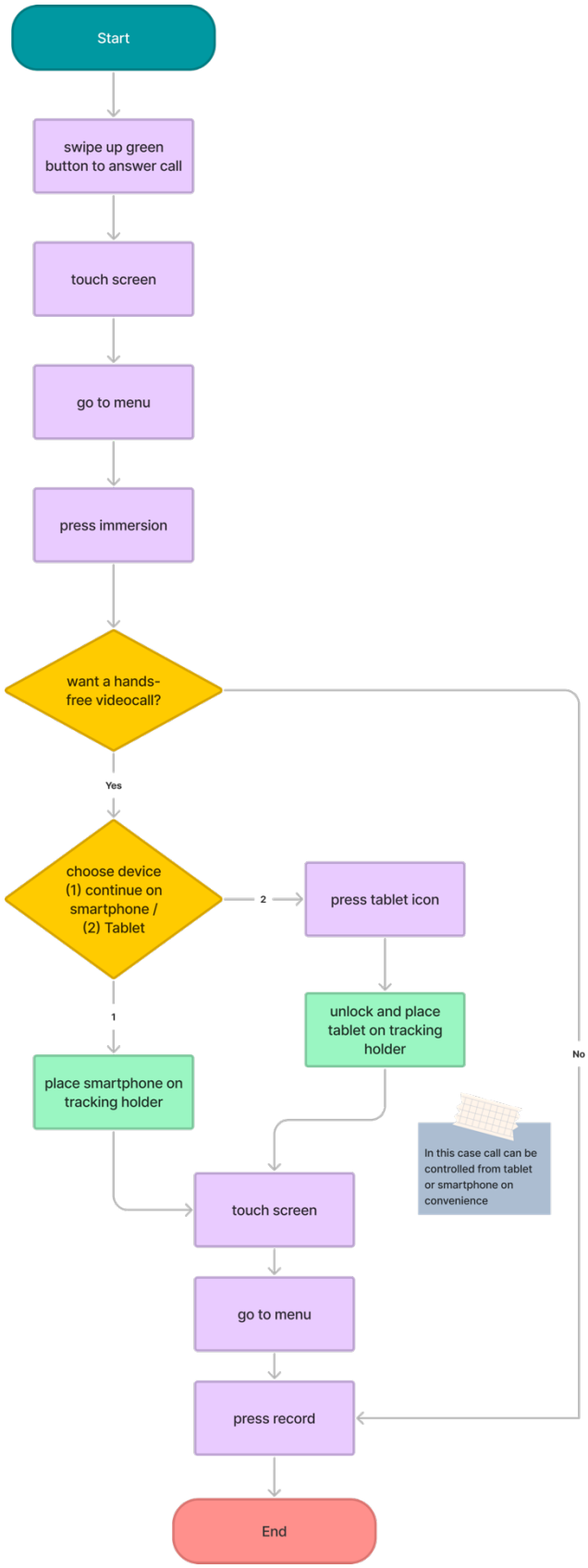
2. Link new device



3. Transfer to TV and use filters



4. Record a call on immersion mode

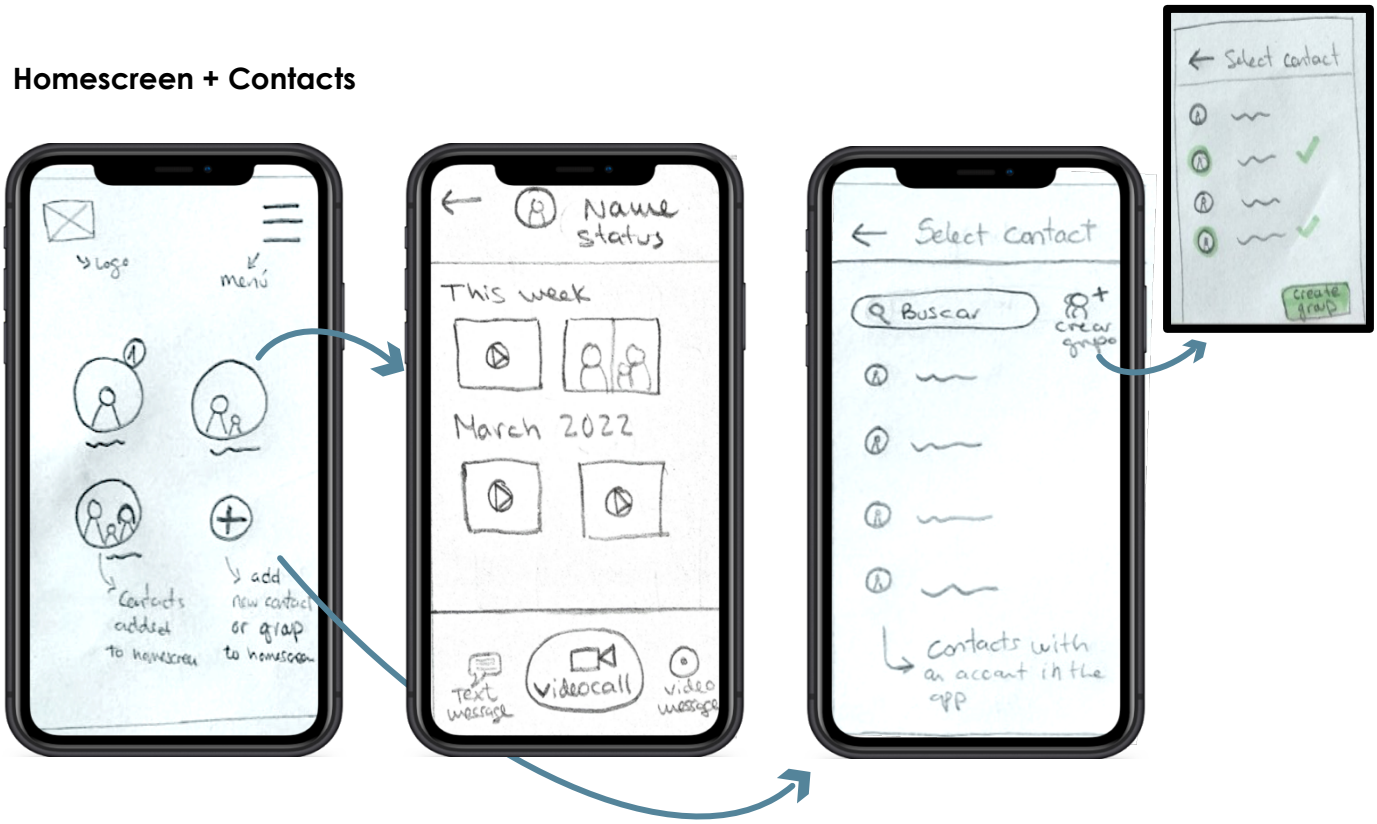


*All interactions start on smartphone.

5.3 Sketches

Once navigation flows are defined, we can start sketching. This technique allows to visualize all elements that conform the appearance of the system (contents, features, etc.) and explore different ideas before making a definitive prototype. Sketch of the physical elements that conform the system can be visualized on Fig. 28, down below the sketches of the main screens of our app:

Homescreen + Contacts



Videocall: transfer to TV/immersion

5.4 Prototype

The high-fidelity prototype of the system has been created on Figma.

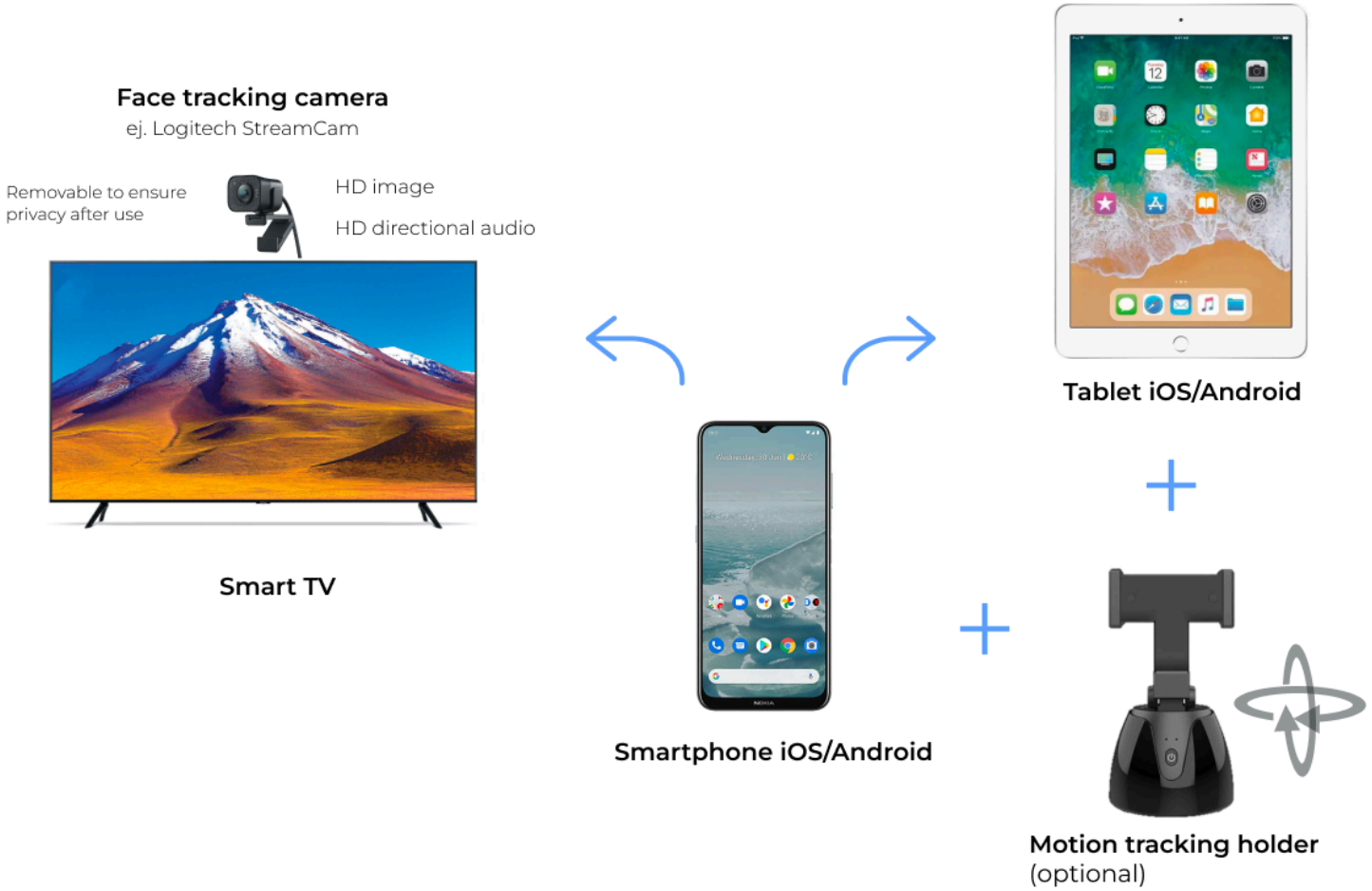
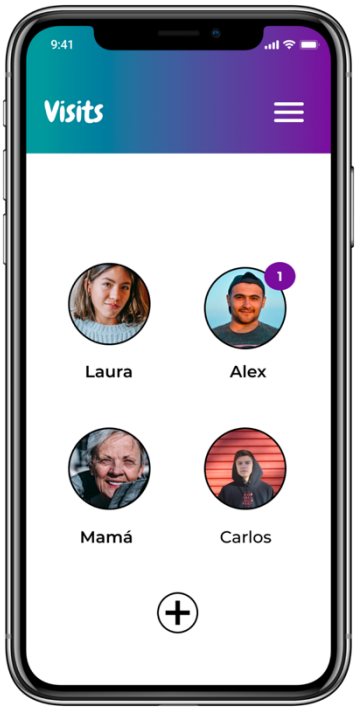
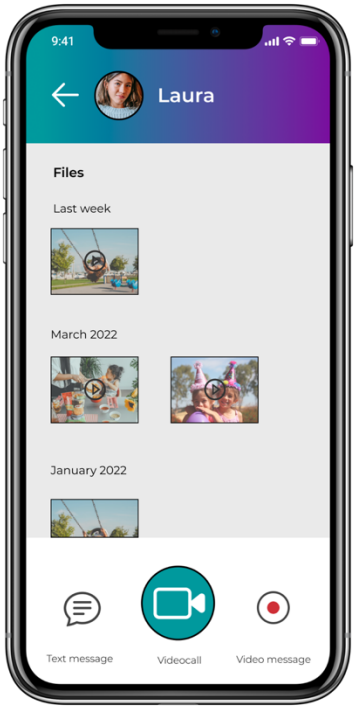


Figure 24. System´s elements (hi-fi)

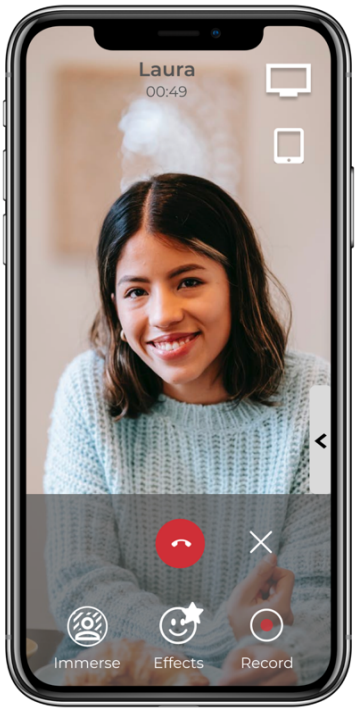


Homescreen



Contact

Calling



Videocall Menu

Immersion mode Smartphone and Tablet

Call transferred to Smart TV (Controls on smartphone)

Improvements from Sketches:

Some details have been changed from the initial sketches for a better interaction:

- In-call menu has been located on the lower part of the screen.
- Choice of language and “front-back camera” feature have been added.
- Onboarding simplification and first steps guide.

5.4.1 Interactive prototype

All screens and interactions can be consulted here:

[Complete prototype]

[Interactive prototype]

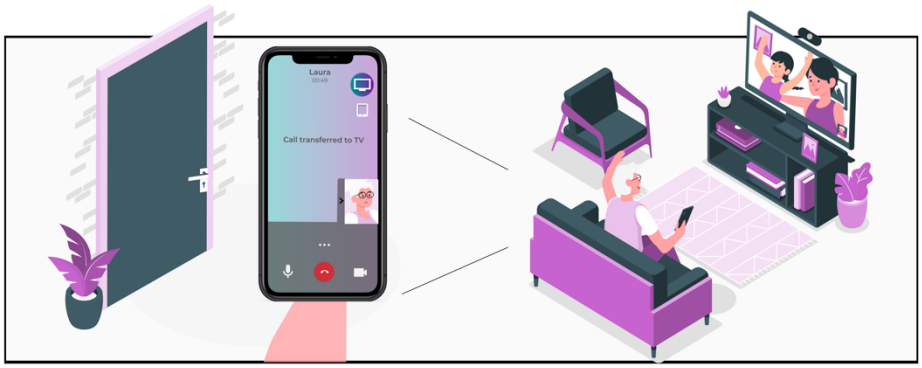
5.5 Storyboard

The purpose of the following storyboard is communicating the interaction with the system and the integration of all its components in a visual way. It will also serve as a tool in the evaluation with users that will be performed next and explained in the next section.

Persona: María

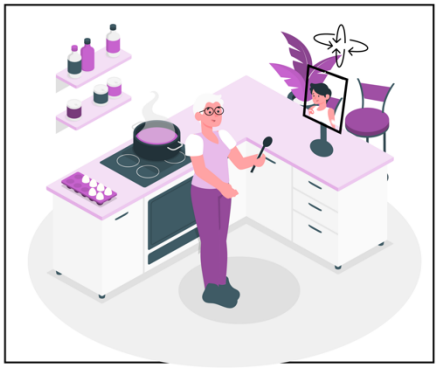
User story: Video communicating with her family in different situations

1. María receives a videocall from her daughter while having a walk in the park.



2. She arrives home and wants to continue the meeting comfortably in her living room. Switches on her TV and transfers the call with one click. She can perfectly see the new dance her grandaughter has learned thanks to the motion tracking camera on the other side.

3. It's almost dinner time and María has to start cooking but wants to tell her daughter about her painting lessons. She mounts her tablet on the holder and transfers the call again. Her daughter suggests her to try the new immersive functionality.



4. She is free to move around the kitchen while talking to her daughter. It feels like having her on the other side of the table!

Figure 25. Storyboard

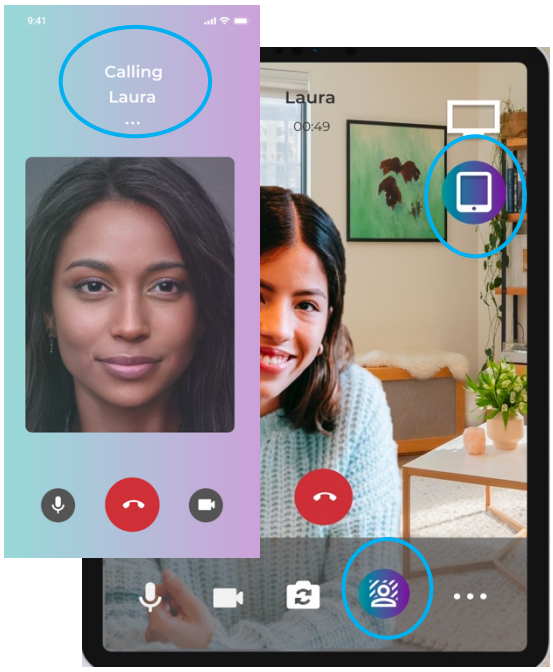
Evaluation

6.1 Heuristic evaluation

Heuristic evaluation is one of the best-known expert methods for usability testing without users. For the analysis of this prototype, I have used the 10 usability rules proposed by Jakob Nielsen (1994). Checking whether our product complies with these heuristics we can detect flaws and strengths. Once the analysis is done, problems identified will be organized by frequency, impact and persistence.

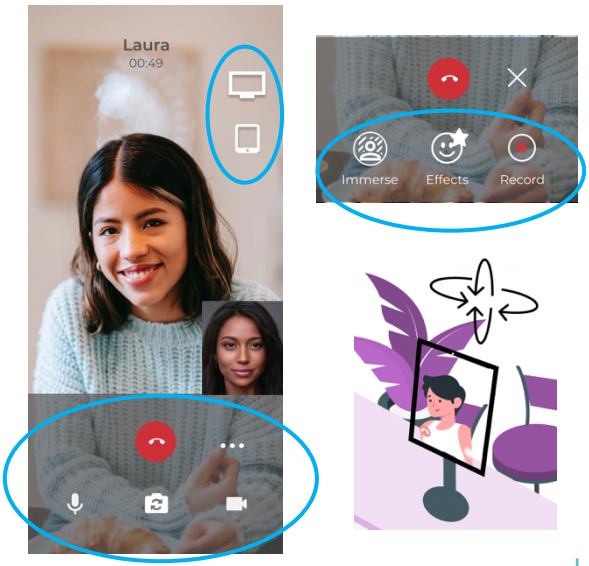
1. Visibility of System status

Icons visibly change background color when they are active and background disappears when they are not. Important actions being performed are explicit: "Calling..., call transferred to...". Call is directly visible on tablet or TV when transferred.



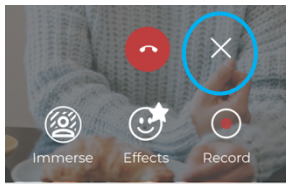
2. Match between system and real world

Most of the icons representing actions are universally recognized, the new immersion feature required a new icon, but all features from the menu are subtitled with their name to make them more obvious. The motion tracking holder makes the tablet interact as a person would do face to face.



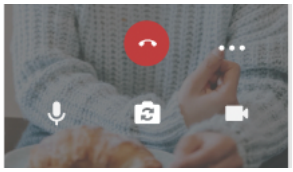
3. User control and freedom

User can always go back to the previous section using the back arrow, close a menu using the X or tapping out of the menu. Features can be deactivated by pressing again on the colored icon (this might be slightly problematic at the beginning) and finish the call at any moment (hang-up icon is visible in all screens). User can remove the camera from TV and choose if using the motion tracking holder depending on the situation.



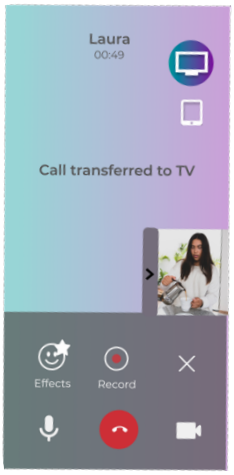
4. Consistency and standards

In-call controls are standard to every VC app. Main menu and In-call menu are represented differently, apart from that all icons representing the same action are consistent throughout the app. Controlling the call on TV with the smartphone can be confusing at the beginning, but on the other hand it is the same concept of a regular tv remote.



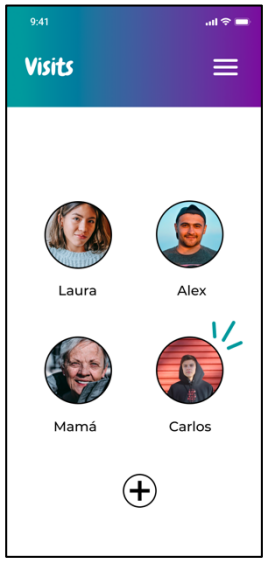
5. Error prevention

Actions/features that cannot be activated in that mode are not visible (ej. Immersion on TV mode), asks for confirmation to link new devices, icons to transfer the call are located top right of the screen (comfortably reachable but not easy enough to activate them by mistake).



6. Recognition rather than recall

Adding favorite contacts to the homescreen makes easy to call them in one step. Subtitles on in-call features also help not having to remember what each icon is for.



7. Flexibility and efficiency of use

The app requires the same level of expertise from everybody, it has been designed to result intuitive even for people with low tech skills. Shortcuts for experienced users don't make a lot of sense in this case.

8. Aesthetic and minimalistic design

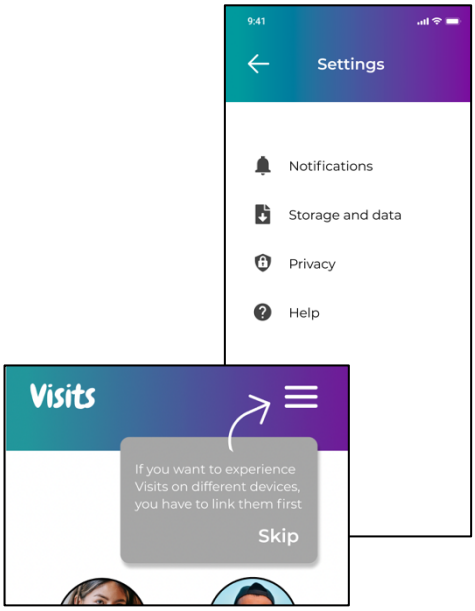
The app has been designed taking into account this principle, a reduced number of options is available in each menu, there is enough space between controls and icons are grouped per type/category.

9. Recognize, diagnose and recover from errors

This aspect was not included in the prototype but it is mentioned in the requirements for the final product: a message will appear when the signal is poor and the video quality will be lowered to preserve audio quality. Possible issues when connecting the motion tracking holder have not been considered.

10. Help & documentation

A Help section is available in the settings section of the main menu as well as a guided onboarding and guide with first steps.



Improvements:

- **Error messages** and possible errors involving the other devices must be included.
- **Informative (first-time) messages** about transferring calls, motion tracking and new immersive functionality they are used might be beneficial.
- **Representation of menus** could be standardized.

6.2 User test

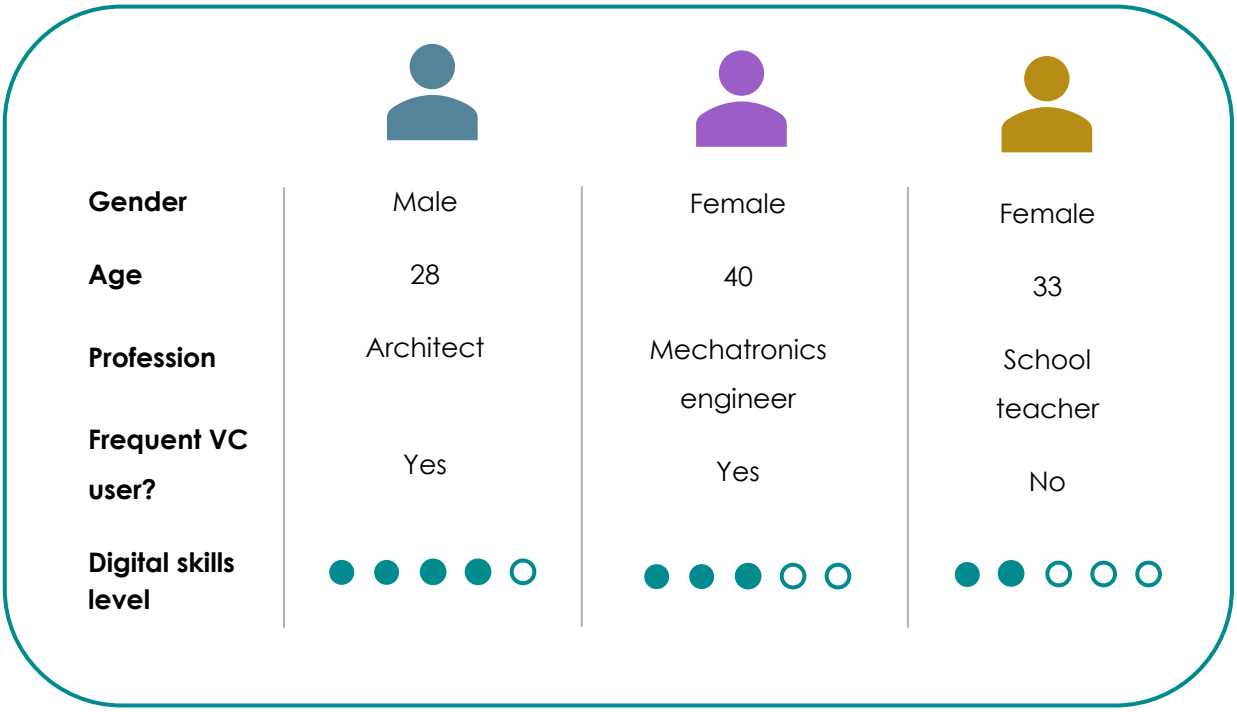
The goal of this evaluations is to test the interface of the app in order to find inconsistencies or usability issues that haven't been identified during the heuristic evaluation. The global concept of the system and way of interacting with it will be evaluated too.

6.2.1 Goals

- Find navigation issues and problems to find features.
- Identify problematic or ambiguous content labeling and presentation.
- Check the actual level of user control and freedom.
- Asses how flexible and intuitive the UI is for users with low tech skills.
- Check user's attitude and opinion towards the system and the interaction globally.

6.2.2 Participants

This first evaluation of the concept and UI was performed with 3 participants. A further evaluation will be performed once the app is developed, adapted to different platforms and, integration (and interaction) with external devices is possible.



Gender	Male	Female	Female
Age	28	40	33
Profession	Architect	Mechatronics engineer	School teacher
Frequent VC user?	Yes	Yes	No
Digital skills level	●●●●○	●●●○○	●●○○○

Table 7. User test participants

6.2.3 Methodology

First part is a **moderated** usability test, 2 of them were online, performed using Zoom for videoconferencing and one of them was in-person, recorded with an external camera. Second part is an **unmoderated "workshop"** based on the collaborative tool Figjam, where graphic material, including the use case storyboard, is presented and participants are asked about their feedback using a template called: "I like, I wish, I wonder".

Pre-test

Once a date and time was agreed, participants received an informed consent (Annex), including information about what type of data will be collected and privacy matters, that they need to return signed. Online participants received the ID for the Zoom call too.
At the beginning of the online meeting, the project will be introduced as well as the procedure of the test. The concept and elements of the system will be explained with the help of some graphical materials developed in the creation section, by sharing screen. Comments and questions expressed in this introductory part will be registered.

After this, participants receive the link to the functional prototype in the chat, they are instructed to open it, share their screen and follow the instructions to perform the required tasks.

Tasks

Task 1: Onboarding.

Scenario: You just downloaded the Visits app and opened it for the first time. You want to set everything up so that you can use it to video call your family that day.

Specific task: Complete the onboarding process, add Carlos to your contacts and link 2 devices.

Success criteria: The onboarding is successfully completed; a new contact is added and two new devices are linked.

Task 2: Record videocall on TV

Scenario: You agreed to connect with Alex that afternoon, but he sent you a message and is not available today. When you get home, you decide to call your sister Laura so she can show you her new apartment. During the call, you want to be comfortable on your couch and have a good

quality image, so you transfer the call to your TV. You want to record it to show it later to your boyfriend/girlfriend/flat-mate.

Specific task: Check Alex's message, call Laura, transfer the call to your smart TV and record it.

Success criteria: Four subtasks have been successfully completed.

Task 3: Immersive mode on tablet

Scenario: After seeing her apartment you realize that your kitchen is a mess and you want to clean it before dinner, but you want to keep chatting with your sister. For that, it transfers the video call to your tablet and places it on the motion tracking stand, so that she can watch it while cleaning and you can also watch it. You also want to test the new immersive functionality.

Specific task: stop recording, transfer the call to your tablet (place your tablet on the stand) and activate the immersion mode.

Success criteria: Three subtasks are completed without major problems (the order is not relevant).

Metrics

- Success or failure per task. How many participants have been able to complete the task, how many did not, how many thought they correctly completed it but didn't (false positive) and how many thought they failed but didn't (false negative)
- Time per task (efficiency)
- Mistakes and deviations from the ideal path
- Perception of complexity. After each task the person evaluates its difficulty (easy – neutral – difficult).
- User's satisfaction. This data with be extracted from comments and post-test questions.

Post-test questionnaire

It will consist of open questions related to the user's interaction with the interface, to dig deeper on difficulties or striking behavior identified.

After the interaction with the prototype, users will be redirected to the Figjam file, and instructed on checking the graphical material and using the template below to share their feedback on post-its.

6.2.4 Results

Online moderated user test

			
Task 1. Onboarding	 1 min 50 sec Easy	 1 min 33 sec Easy	 1 min 20 sec Easy
Task 2. Record Videocall on TV	 51 sec Easy	 1 min 13 sec Neutral	 1 min 24 sec Neutral
Task 3. Immersive mode on tablet	 30 sec Easy	 14 sec Easy	 58 sec Easy

Table 8. Results user test

Task 1. Onboarding

Participant 1: Completed. Participant got distracted with a message alert, but that did not affect completing the task.

Participant 2: Completed – False Negative. Participant went directly to settings to link new devices, missing the right option, but she went back to home screen and found the option easily on the second attempt. She linked the two devices correctly but got confused by the "status" of the device and tried again.

Participant 3: Completed – False Negative. First part of the task went smooth but, same as participant 2, she linked the new devices successfully but got confused by the status of the device and thought she was not doing it correctly.

The three participants perceived the task as **easy.**

Task 2. Record videocall on TV

Participant 1: Completed, but it was observed that the participant tried to close the message after reading it, tapping on the white space, and was not able to do it.

Participant 2: Completed with issues. Transferring call to TV was easy for her but she struggled with finding the menu on the smartphone when the call was transferred. She expressed that it took some time to understand that what she was seeing was a representation of the smartphone but that, if she would have had it for real on the hand "she would have found it easily" because it is pretty clear. That indicates that the problem was the way of representing it rather than the prototype itself.

Participant 3: Completed with issues. Same as previous participant, first part of the task was ok, but she struggled finding the "record" button when the call was transferred: "My attention was on the TV... I think if I had the smartphone in my hand, it would be more obvious but doing it on my laptop it wasn't so clear".

One participant considered the task **easy,** two considered it **neutral.**

Task 3. Immersive mode on tablet

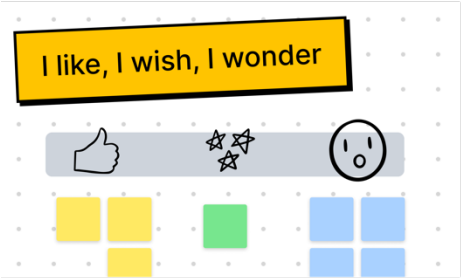
Participant 1: Completed. In task 2, the participant made a couple of extra steps (finished the call) so, this task was started in a different point than the originally planned, but could be completed successfully. This indicates that the prototype is flexible and offers different paths to do the same action.

Participant 2: Completed. Three steps completed without issues.

Participant 3: Completed. There was some confusion on the type of device where the call was being sent so it required a bit of exploration from the participant, but she eventually performed the task completely.

All participants perceived the task as **easy**.

Unmoderated whiteboard session



After seeing the graphical content: structure of the system and interaction storyboard; participants were asked to write their thoughts on this template (including comments about the tasks performed if they didn't verbalize them before).

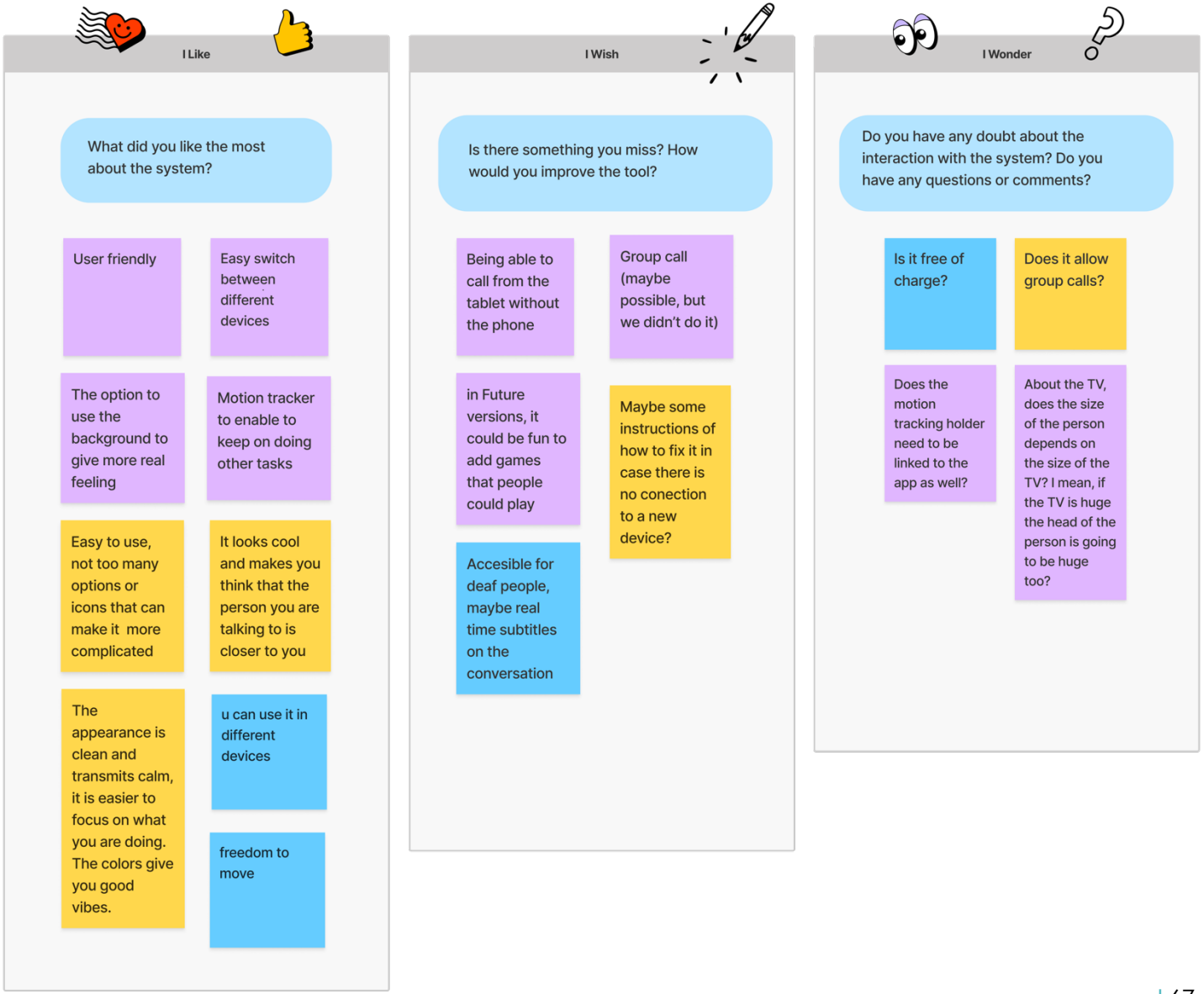


Figure 26. Figjam whiteboard with feedback

Other questions/comments registered:

"Can it be used on the tablet independently?"

"I think it was very easy, the only issue I had (recording call on TV) is because I did not realize that what I was seeing was the smartphone in my hand"

"About having the call on the TV, it would be like casting with Netflix right?"

"How do you know the motion tracker is ON?"

Limitations:

Tough our participants differ on affinity with technology and digital skills; we could not get a varied sample in terms of age. Therefore, to confirm that the tool is intuitive for all kind of users, it is advised to count with people from a younger (<20) and older age (>50) in a future user test.

Real interaction with the different devices was not possible yet, all elements were represented in Figma.

6.2.5 Insights and improvements

The concept and the prototype were in general positively evaluated and participants could complete all the tasks. What participants **liked** the most:

- **Easy to use**
- **Flexible: adapted to different devices**
- **Motion tracking**
- **Freedom of movement**
- **Immersive option**
- **Minimalistic and aesthetically pleasant**

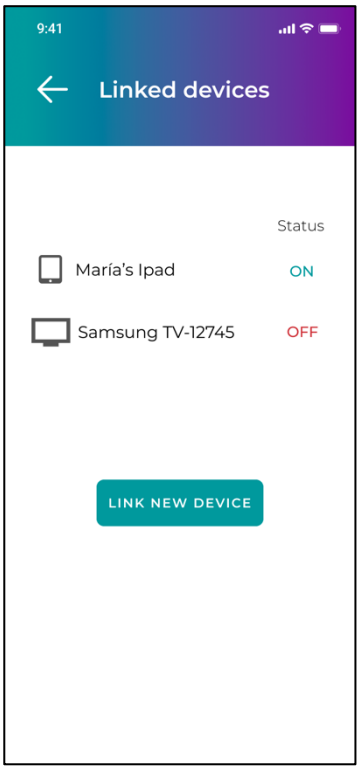
Areas of **improvement** identified in order of importance:

- **Status of linked devices can be confusing.** Possible solutions: add a confirmation message when devices are successfully linked. Instead of the status of the system always there, show error message only when trying to transfer a call and device is disconnected.

- **Representation:** for Future tests, more clear representation of smartphone screen to control the call when transferred to TV, that enhances visibility.
- **Help/error messages** if linking a new device fails. This matches one of the flaws identified during the heuristic evaluation.
- **Design for group calls**. Tough, best results can be achieved on one-to-one calls, users want to be able to make small group calls so that option should be represented as well.
- **Accessibility:** add subtitles for deaf people.
- **Closing text messages:** Allow closing pop-up messages by tapping on the white area
- **Notifications when a device is connected**
- **Adjust image dimensions,** so the size of the person always looks realistic, no matter the size of the TV.

6.3 Changes implemented

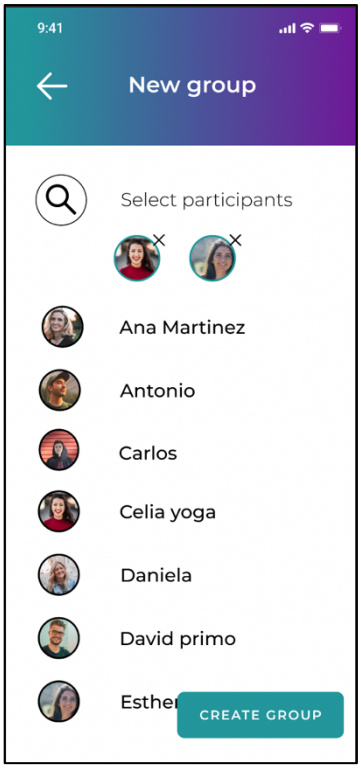
Areas of improvement identified during the test with users and heuristic evaluation have been combined and led a design iteration to improve the tool with the following changes:

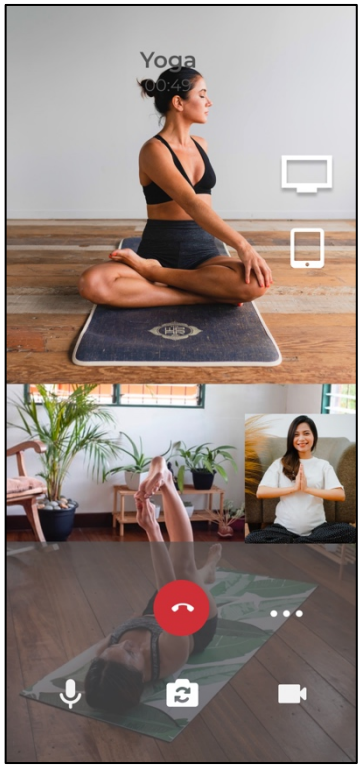


1. Status of Linked devices and error message. The status of the linked devices has been modified to make it more clear and avoid confusion. If a device is not connected or out of network, the icon appears less salient and an error message will show up when trying to transfer the call to it.

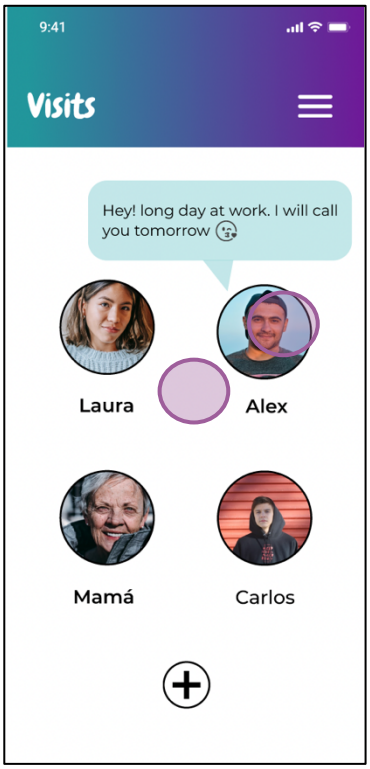
2. Representation of the call transferred to TV has been improved to make smartphone controls more salient by increasing the size of the smartphone and adding a hand holding in in front of the TV to make the interaction more obvious.

3. Accessibility. An option to activate real time subtitles has been added to the menu for deaf or hard of hearing users.





4. Design for group calls. The prototype has been extended making possible to create a group, perform a group call and transfer it to other devices, including the use of the immersion mode.



5. Closing text message can be done by tapping on contact picture or white area.

6. Notification messages have been added when a device like the motion tracking holder is connected. (When disconnected a message will show up as well as in the example 1).

Final interactive prototype

Conclusions & Future work

7.1 Discussion

Online communication is a challenging design topic that involves many elements to ensure a satisfactory experience. A good usability is basic but not enough. When we talk about personal communication, emotions and enjoyment play a big role as well. That´s why a thorough research, especially research with users has been key to uncover needs and desires, and guide the design of this tool. An iterative process that facilitates continuous improvement has been also very important. Also, exploring new technologies, like augmented reality, has served to provide with an innovative solution that users seem to enjoy, and on a personal level, I could learn a lot about the potential and limitations of this technology.

Integrating my work as an intern at TNO (Human factors research) in a design plan like this hasn't been easy, but I think I got a more complete and richer project as a result.

Lastly, designing a flexible multiscreen product, for different situations and needs, has also been a challenge, since the specific characteristics of each device and context of use need to be considered and combined in an effective way.

7.2 Future work

On a more technical level, further development of this tool should take into account its implementation on different platforms, an effective image resizing depending on the dimensions of the device used (TV) and improve the motion/voice tracking as much as possible.

Regarding UX, I encourage to continue exploring solutions to address some of the pain points that could not be covered in this project and keep making distance feel smaller, like ways of achieving a more realistic eye-contact, the implementation of innovative XR devices or the addition of some kind of haptic or olfactive experience. For this more research is needed.

Also, it is recommended to test the tool again, with a bigger and more varied sample of users, once it is developed and a real videocall can be made in different scenarios integrating the three devices.

7.3 Conclusion

This project had some limitations in terms of time and resources, for example, ideally the proof-of-concept would have included other devices to test, like AR glasses and the creation phase would probably be richer if a multidisciplinary team would be involved; but, all in all, I think the goals have been achieved and the outcome is a valuable product:

- Trends, other solutions and user's needs on personal communication were thoroughly explored.
- An experiment testing the benefits of augmented reality on VC was performed.
- Definition of requirements, based on the previous, guided the design of the new tool and interaction design principles were followed to come up with a high-fidelity prototype that was ultimately tested with users.

As a result, Visits is an intuitive tool that makes personal video-communication a more natural and immersive experience.

Glossary

English

Augmented reality (AR): a technology that superimposes a, usually computer-generated, image on a user's view of the real world, thus providing a composite view.

Extended reality (XR): referring to all real-and-virtual combined environments and human-machine interactions generated by computer technology and wearables

Social presence: social presence theory explores how the "sense of being with another" is influenced by digital interfaces in human-computer interactions.

Telepresence: the use of technologies that allow remote control of machinery or apparent participation in distant events.

Usability: the capacity of a system to provide a condition for its users to perform tasks safely, effectively, and efficiently while enjoying the experience.

User Centered Design: iterative design process in which designers focus on the users and their needs in each phase of the design process.

Virtual visit: term used in the scope of this project to denominate a personal online meeting.

Video mediated communication: interpersonal interaction via the use of computers or other digital media featuring video and audio signals.

Zoom fatigue: exhaustion experienced from using any kind of video calling interface to have virtual interactions with others.

Español

Realidad aumentada (RA): tecnología que superpone una imagen, usualmente generada por ordenador, sobre la visión que el usuario tiene del mundo real, creando una imagen compuesta.

Realidad extendida (XR): se refiere a todos los entornos que combinan lo real y lo virtual e interacciones humano-máquina generadas por ordenador que implican el uso de *wearables*.

Presencia social: la teoría de presencia social explora cómo "la sensación de estar con otra persona es influenciada por interfaces digitales en el ámbito de la interacción humano-ordenador.

Tele-presencia: se refiere al uso de tecnologías que permiten el control remoto de maquinaria/robots o la participación en eventos remotos.

Usabilidad: la capacidad de un sistema para proporcionar la condición óptima para que los usuarios realicen sus tareas de forma segura, efectiva y eficiente, disfrutando la experiencia.

Diseño centrado en el usuario: proceso de diseño iterativo donde los diseñadores se centran en los usuarios y sus necesidades en cada fase del proceso.

***Virtual visit*:** término usado en este proyecto para referirse a las video-llamadas de carácter personal/doméstico.

Comunicación mediada por video: interacción interpersonal a través del uso de ordenador u otro medio digital usando señales de video y audio.

***Zoom fatigue*:** fatiga experimentada a causa del uso de cualquier tipo de interfaz de video-llamada para tener interacciones virtuales.

Bibliography

Ahn, D., Seo, Y., Kim, M., Kwon, J. H., Jung, Y., Ahn, J., & Lee, D. (2014). The effects of actual human size display and stereoscopic presentation on users' sense of being together with and of psychological immersion in a virtual character. Cyberpsychology, behaviour and social networking, 17(7), 483–487. https://doi.org/10.1089/cyber.2013.0455

Apostolopoulos, J. G., Chou, P. A., Culbertson, B., Kalker, T., Trott, M. D., & Wee, S. (2012). The road to immersive communication. Proceedings of the IEEE, 100(4), 974-990.

Biocca, F., Burgoon, J., Harms, C. & Stoner, G. (2001). Criteria And Scope Conditions For A Theory And Measure Of Social Presence.

Biocca, F. Harms, C. and Burgoon, J.K. (2003). Towards a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. Presence, 12(5), 456-480.

Blanche, P. A. (2021). Holography, and the future of 3D display. *Light: Advanced Manufacturing*, *2* (4), 1-14.

Choi, M., & Choung, H. (2021). Mediated communication matters during the COVID-19 pandemic: The use of interpersonal and masspersonal media and psychological well-being. *Journal of Social and Personal Relationships*, *38*(8), 2397–2418. https://doi.org/10.1177/02654075211029378

Fauville, G., Luo M., Queiroz, A.C.M., Bailenson, J.N., Hancock, J. (2021). Zoom Exhaustion & Fatigue Scale, Computers in Human Behavior Reports, Volume 4, 100119, ISSN 2451-9588. https://doi.org/10.1016/j.chbr.2021.100119.

Gooch, D., & Watts, L.A. (2015). The Impact of Social Presence on Feelings of Closeness in Personal Relationships. *Interact. Comput., 27*, 661-674.

Gunkel, S. N., Stokking, H. M., de Koninck, T. J. I., & Niamut, O. A. (2019). Everyday Photo-Realistic Social VR: Communicatie and Collaborate with an Enhanced Co-Presence and Immersion.

Hauber, J., Regenbrecht, H., Hills, A., Cockburn, A., & Billinghurst, M. (2005). Social presence in two-and three-dimensional videoconferencing.

Huisman, G., Frederiks, A. D., & Heylen, D. (2013, September). Affective touch at a distance. In 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (pp. 701-702). IEEE.

Ishida, T., Sakuraba, A. & Shibata, Y., (2011).“Proposal of high realistic sensation system using the large scale tiled display environment,” In Proceedings of Network-Based Information Systems (NBiS), 444–449.

Kirk, D. S., Sellen, A ., & Cao, X. (2010, February). Home video communication: mediating'closeness'. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work* (pp. 135-144).

Koh, E. (2010). “Conferencing room for telepresence with remote participants,” In Proceedings of the 16th ACM international conference on Supporting group work, GROUP’10, ACM, 309–310

Li, J., Kong, Y., Röggla, T., De Simone, F., Ananthanarayan, S., De Ridder, H., & Cesar, P. (2019, May). Measuring and understanding photo sharing experiences in social virtual reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-14).

Miller, M., (2020). # AloneTogether–An Exploration of Social Connectedness Through Communication Technology During Physical Distancing.

Nielsen, J. (1994). 10 Usability Heuristics for User Interface Design. Retrieved 12 May 2022, from https://www.nngroup.com/articles/ten-usability-heuristics/

Project Starline: Feel like you're there, together. (2021). Retrieved 23 March 2022, from https://blog.google/technology/research/project-starline/

Skalski, P., Tamborini, R., Shelton, A., Buncher, M., & Lindmark, P. (2011). Mapping the road to fun: Natural video game controllers, presence, and game enjoyment. New Media & Society, 13(2), 224–242. https://doi.org/10.1177/1461444810370949

Techopedia.com. 2017. *What is Telepresence? - Definition from Techopedia*. [online] Available at: https://www.techopedia.com/definition/14600/telepresence
[Accessed 3 March 2022].

UOC Design Toolkit | Test con usuarios guía. (2022). Retrieved 5 May 2022, from http://design-toolkit.recursos.uoc.edu/es/guia/test-con-usuarios/

Yu, K., Gorbachev, G., Eck, U., Pankratz, F., Navab, N., & Roth, D. (2021). Avatars for Teleconsultation: Effects of Avatar Embodiment Techniques on User Perception in 3D Asymmetric Telepresence. *IEEE Transactions on Visualization and Computer Graphics*, *27*(11), 4129-4139.

Annex

INFORMED CONSENT

Interview about videoconferencing experiences

This interview is part of a research project about modern video-communication and is performed by the TNO independent research institute (www.tno.nl). Through this interview we want to investigate to what extent currently available video-communication systems fulfill the need for social communication between people in affective long-distance relationships, and what aspects of these systems need further improvement. Your participation will help us to get valuable insights, that we can use in the future for the development of new communication systems. During an online interview a TNO researcher will ask you questions about your experience using video communication systems and your opinion on different aspects of the same topic. During the interview, you will not be asked about any sensitive information and you do not need to answer any questions you do not feel comfortable with.

The estimated length of the interview is 20 min. You can take a break or quit the interview at any moment if you need it, without the need of giving further explanation.

After the interview, you will be asked to complete an online questionnaire where you have to rate some statements according to your experience with video-calling. The estimated length of this questionnaire is 5 minutes.

After the completion of the study (interview + online questionnaire), you will receive a voucher of 10 euro at bol.com.

Your answers will be completely anonymous, only your age and gender will be registered for demographic purposes. The content of the interview (audio) will be recorded for further analysis. These recordings will be treated as confidential and will not be shared outside the scope of this project.

You are eligible to participate in the experiment if...

- You are between 18 and 65 years old

- You are fluent in the English language
- You have a long-distance affective relationship (of any kind)
- You frequently use video conference tools to communicate with the other person/s in that relationship

You are not eligible to participate in the experiment if…

- You have vision or auditory problems (after correction methods).
- You participated in earlier Social XR studies.

For any further questions about the study, please contact:

Marina Álvarez

Email: marina.alvarezmerida@tno.nl

Undersigned,

Name _____

Date of birth _____

declares to voluntarily participate in this study, entitled

Ervaringen met moderne videocommunicatiesystemen

(Experience with modern video-communication systems)

I confirm that I have read the information about the above research and I understand the information. ☐

The intention of the interview has been explained to my satisfaction. ☐

I know that my participation in the study is completely voluntary and that I can withdraw my consent at any time without having to give a reason. ☐

I give permission for my data to be stored and process by authorized members of the investigation team, for the purposes described in the information. ☐

I give permission to reuse my research data for future research in the research area described, provided that it is coded in such a way that it can no longer be traced back to me as a person.
☐

Place and date _____

Signature _____

INTERVIEW SCRIPT

1. How often do you use videocalls to communicate with your loved ones? With whom do you videocall most often?

2. For what kind of conversations/topics is videocall the preferred communication method with your remote loved ones? Why?

3. Which device do you use most frequently for your personal videocalls? Smartphone/Tablet/Laptop. Why?

4. What's your preferred app for personal videocalls? Why?

5. Which place do you usually choose to have your personal videocalls? Why? How does this environment look like?

6. What do you miss when communicating via videoconference vs. face to face?

7. Have you experienced issues or are there things you don't like from videoconferencing apps you frequently use? What??

8. How has video calling influenced your relationship with loved one/s that live far from you?

9. How do you envision the future of remote communication? In like 10 years

10. Imagine there is a new state-of-the-art communication system on the market. What is the characteristic that would make you try it?

11. How likely would you choose a tool that allows you to really feel "as if you were in the same room" with the other person, over regular videoconferencing apps? 1 to 10 Why?

12. How do you feel about using modern technologies that are still not familiar to you for remote communication?

13. Have you ever experienced AR/VR? (explanation if needed) Do you think this technology can contribute to the improvement of daily communication systems? How?

14. Would you have any concern about using this kind of technology for your private communication?

15. Do you have any comments or suggestions?

QUESTIONNAIRE

Videoconferencing experiences

This study is part of a research project about modern mediated communication technologies from the independent research institute TNO (www.tno.nl). The goal of this questionnaire is to get information about your experience with current videoconferencing systems for personal communication (not work-related). Your participation will help us to get valuable insights in order to improve such systems.

Your answers will be completely anonymous. Your data will be stored and used only within the scope of this project. *

- [ ] I give permission for my data to be stored and processed by authorized members of the investigation tea...

Gender *

- ( ) Male
- ( ) Female
- ( ) Prefer not to say

Age *

- ( ) 18-25
- ( ) 25-35
- ( ) 35-45
- ( ) 45-55
- ( ) 55-65
- ( ) 65+

How frequently do you use videoconferencing systems for personal use (with video)? *

- ○ Several times a day
- ○ Once a day
- ○ Several times a week
- ○ Once a week
- ○ Less frequently

What is your preferred tool for personal videocalls? *

Texto de respuesta corta

To what extent are you satisfied with your most used videoconferencing tool?

	1	2	3	4	5	
Not at all satisfied	○	○	○	○	○	Completely satisfied

How often your personal videocalls are group calls? (more than 2 devices involved) *

- ○ Never
- ○ Rarely
- ○ Sometimes
- ○ Often
- ○ Always

How often do you perform activities using a videoconferencing tool? (like sports, workshops, games, etc). *

- ○ Never
- ○ Rarely
- ○ Sometimes
- ○ Often
- ○ Always

If so, which ones?

Texto de respuesta corta

When having personal videocalls, how important is it for you... *

	1. Not at all important	2	3	4	5. Extremely important
to have the impression that the other person and you are together in the same room?	○	○	○	○	○
that the conversation with the other person feels natural (turn taking, eye contact, etc.)?	○	○	○	○	○

to see the other person's expression and understand his/her emotions?	○	○	○	○	○
being able to move around your own environment / room?	○	○	○	○	○

During COVID crisis, have you used videoconferencing systems (for personal use) *

- ○ More than before
- ○ Same as before
- ○ Less than before

Right after a long (private) videoconference: *

	1. Not at all	2. Slightly	3. Moderately	4. Very	5. Extremely
How tired do you feel?	○	○	○	○	○
How mentally drained do you feel?	○	○	○	○	○
How much do you need time just by yourself?	○	○	○	○	○

How much do you tend to avoid social situations?	○	○	○	○	○
How much do you feel like doing nothing?	○	○	○	○	○
How much do you feel too tired for doing other activities?	○	○	○	○	○
How emotionally drained do you feel?	○	○	○	○	○
How irritable do you feel?	○	○	○	○	○

Would you purchase a product which only purpose is to provide you with improved personal videocalls? (better quality / innovative features / more immersive) *

- [ ] Yes
- [ ] No

Would you dedicate a space at home for this product and purpose? *

- [ ] Yes
- [ ] No

INTERVIEW TRANSCRIPTION EXAMPLE

1st Interview.

Gender: Female

Age: 28

Nationality: French

Occupation: Designer

16. **How often do you use videocalls to communicate with your loved ones?** Once or twice a week

17. **With whom do you videocall most often?** With my family, parents brother and sisters. My sister got a baby so it is nice to see him growing. Or with another friend I work with, she also got a baby and she uses the video to show me her work or the baby... it is very much related to kids (The video).

- Most of the times with people that is very close to you? Yes.

18. **For what kind of conversations/topics is videocall the preferred communication method with your remote loved ones? Why?** More for fun, for example when I call with my family sometimes we end up putting stupid filters... it's really random, just to know how we are, show the weather. I had a professional call with someone and it was nice to have the video as a first encounter to get a bit closer relationship and also it was about showing space.

- To see the environment of the other person? Yes

19. **Which device do you use most frequently for your personal videocalls? Smartphone/Tablet/Laptop. Why?** Phone because it is easier, I have unlimited internet and good quality, when I do it from my laptop I am tied to how fast the wifi goes. It is portable, for the filters, fun, flexible to move around and show space from different angles.

20. **What's your preferred app for personal videocalls? Why?** WhatsApp, works better, quality, sometimes we use fb messenger for the filters.

21. **Which place do you usually choose to have your personal videocalls? Why? How does this environment look like?** Depends on the time, during the day it is in my studio where I spend most of my day, that would be behind my desk, if it is in the evening it would be at home on my couch, and sometimes maybe walking around. Or

sometimes when I am on my bike… I did it a few times, so the times goes faster and I can show how the weather is.

- So, to show your environment as well? Yes.

22. **What do you miss when communicating via videoconference in comparison with face to face?** I miss seeing the environment of the other person, what she is doing, context, the "hidden information", body language, reactions, nonverbal information. Or for example, my sister cooks a lot and even when she shows me what she is cooking, I cannot smell it, so… you are not so much part of the action.

23. **How has video calling influenced your relationship with loved one/s that live far from you?** Stronger, it makes it easier to have a quick call, less formal, straight to the point… sometimes calling without the video is such a pain… It depends, for example when my dad calls he talks and talks and I just leave the phone and start doing something else (he doesn't see that she is not interested and doing other things). Well, what was the question again? Hahah
    Oh yes, it has made easier to communicate… and I think it also help during COVID, for example this trend of having an Aperitivo with your friends or family together.
    - But did you continue doing it afterwards or was it more like a moment thing? No, I did it 2 or 3 times, then it was more informal.

24. **How do you envision the future of remote communication?** In 10 years
    I like the idea of being able to walk with the person. There was an app or something that develops the idea of VR but like associating the movement with the other person, so you walk together with another person you VR character is also walking. You mean in your real environment? How does it look like? Well… [interruption] I would like to have more of doing things together, like having a coffee break together, walk together to your coffee machine…
    - More like seeing the person in your own environment? Well.. that you can choose depending on the situation.

25. **Imagine there is a new state-of-the-art communication system on the market. What is the characteristic that would make you try it?** It would be curiosity… to feel, to smell to touch better the environment of the other person. Like for example when you have couples in long distance and they make this kind of connected underwear… I see this and I think oh! it looks like a pain… but then if it is easier to use, that you don't have all of this electronic stuff.
    - That you don't feel the technology in between? Yes… like when you have fantastic movies, you see a device that 3d scan the environment and you can immerse in it, you can move around and touch things, you turn and you see what is in the back, you don't need to wear this heavy VR glasses (HMD), the environment is recreated around yourself you can feel it and smell it.

26. **How likely would you choose a tool that allows you to feel "as if you were in the same room" with the other person, over regular videoconferencing apps? 1 to 10** 8 that would be nice **Do you think that would make the experience of conversating remotely more enjoyable?** Yes, It would break boundaries, it would erase maybe the need to travel. You don't need to pay 600 euros to go to the other side of the planet to have drinks with your friend and it would be more ecofriendly as well. If it works it would be a good substitute of ftf meetings. But the quality for now its not so good, there's is always a technical problem... and when there is even a single problem it ca ruin the whole meeting.

27. **How do you feel about using modern technologies that are still not familiar to you for remote communication?** I think I would not try it myself first... I need to have feedback first from other people that I trust, relatives or friends. I am not the kind of person that tries all the new things in the market first.

28. **Have you ever experienced AR/VR?** (explanation if needed) Do you think this technology can contribute to the improvement of communication systems? How? Yes. Yes, in terms of idea yes... it would be fascinating to see a 360 image around you, to be immersed in a virtual environment, experience a different reality. But I think it is a one time thing, I have seen it always at exhibitions or... always related to fun, I would not see myself putting the headset on for a normal personal conversation, it is too complicated. If I don't have to use a headset, I might use it on a more regular base.

29. **Would you have any concern or doubt about using this kind of technology for personal remote communication?** (Ethical, complexity...) maybe there is this idea of the environment you are going to be involved in... in terms of data. What if all this info about your environment would be collected by the app? Like how your environment looks like, what kind of drinks you have, every detail... you would not know what and how much they kno about you. Data privacy is my main concern.

30. **Do you have any comments or suggestions?**
    OVR, platform with avatars cartoons, where you meet other people in virtual representation of famous places like the sixtine chapel. You can choose all the characteristics of the avatar. It was super interesting to be in that environment, it was also a conversation starter with other people. Do you think it would be different if the avatar would be really you? I don't think it changes a lot, because even if you choose your own characteristics for the cartoon... is it really you? There's is always this bridge of not being able to grasp the identity of the other person. You know it is not a natural situation... like when you play Fifa (videogame) or you go to madame Touseao... It looks like the person but it is not, and that is always weird. I'd rather have something completely different than something that looks like a fake replica.

Method

Participants and procedure

Eighteen pairs of participants (*N* = 36) were included in this study [18 Female (50%), *Mage* = 30.78, *SDage* = 6.4)]. All pairs shared a relationship; the highest proportion being close colleagues or classmates (47.2%), followed by romantic couples (27.8%) and friends (19.5%). Recruitment was done through physical and digital posts and most participants were TNO employees, partners and friends of those.

Each pair of participants was randomly assigned to one condition only (between subjects). The experiment consists of two conditions: (1) a regular videoconferencing tool / (2) the TNO AR-based communication tool. The experiment took place in a controlled environment. The day of the experiment, the two participants were placed in different rooms and introduced to the system they will use to interact with the other participant. To stimulate conversation and keep it somehow similar between conditions, we suggested a conversational task: "*Plan your next trip together*" (Place, budget, transport, attractions, etc.). This allows participants to focus on the experience with the tool since it does not require them to pay a lot of attention to any additional task-related items (e.g., written instructions, images, objects etc.) that are not relevant for the experience itself. Participants are free to talk about something else and use the language they naturally would use with the other person. They can finish the conversation at any moment, with a limit of 15 min.

After the conversation, they are asked to answer a questionnaire about the experience. There is currently no standard evaluation methodology to assess UX of AR based social communication (Lee, L. N. et al., 2019). Therefore, in this study we designed an evaluation methodology based on multiple existing questionnaires, objective measure (behavior) and some questions included by ourselves:

Social Presence

- H-MSC-Q of which only the social presence scale will be used (Toet at al. 2021): 10 items
- Networked minds questionnaire (Biocca et al., 2001): 34 items

Interpersonal closeness

- Other in Self (IOS) scale (Aron et al., 1992): 1 item

Overall UX

- The User Experience Questionnaire (UEQ) adapted: 18 items. Due to the specific characteristic of the AR-based system prototype (It still does not have a functional user interface that the user can manipulate independently). Only the Attractiveness, Stimulation, Novelty and Dependability scales from this questionnaire are used. Two of the pragmatic scales (Efficiency and Perspicuity) are not applicable yet.

Quality

- Score Overall quality and image quality (1-7)

Attitude towards AR for VC

- Intention to use (score) and open questions

Behavioral patterns

- For this purpose, the conversations were recorded (participants were previously informed).

Technical setup

The rooms were decorated trying to recreate a living room ambient, to provide a more intimate and realistic atmosphere, since in prior exploratory research (interviews and questionnaires) it has proved to be the most typical space for personal videocalls. In every condition, setup and environment are symmetrical for both participants. In both conditions, audio is recorded by the microphone integrated in the front cameras. We chose a large screen (46 inches) to provide a human-size image. Many previous studies have suggested that a life-sized view is likely to enhance the user's sense of social presence during a videoconference (Ahn et al., 2014; Ishida et al. 2011; Koh, 2010).

Condition 1: Regular videocall (baseline)

An overview of the setup is shown in Fig.20. The screen will be mounted vertically on a stand, which will be positioned on the opposite side of a table from the participant. The participant will be recorded using a Logitech webcam attached to the top of the screen. The videocall will take place using Microsoft Teams software. Both participants will see a 2D image of the other participant and his/her background. Self-view will be disabled to avoid distractions and maintain both conditions similar (the AR condition does not have this feature).

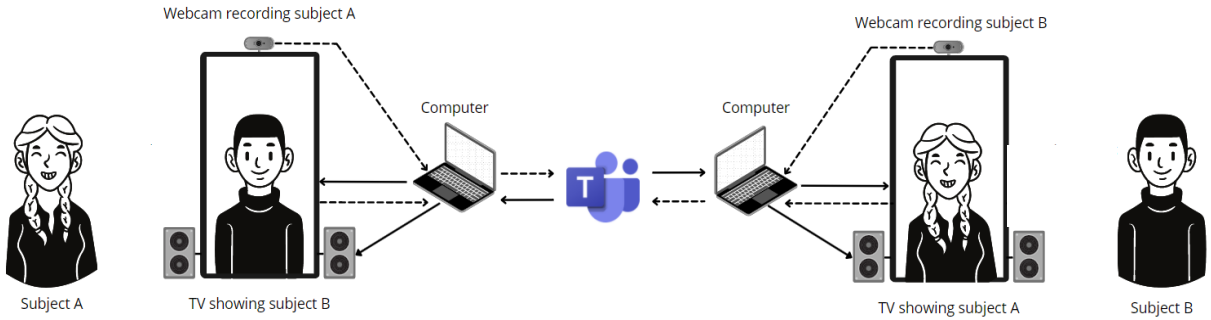


Figure 28. Technical setup condition I

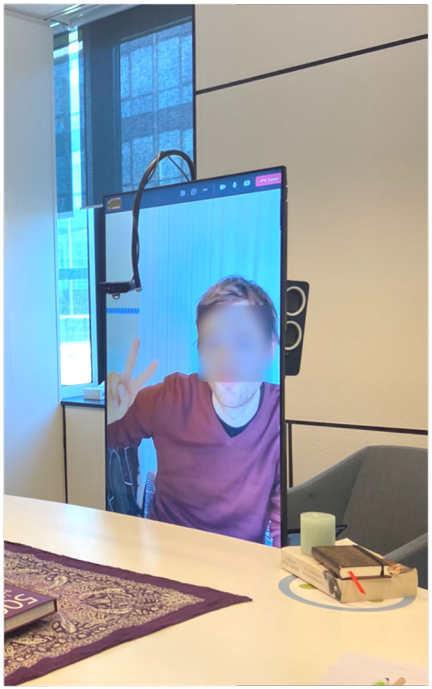


Figure 27. Setup using MS Teams

Condition 2: AR based videocall

An overview of the setup is shown in Fig.21. The screen will be mounted vertically on a stand, which will be positioned in front of a chair on the opposite side of a table from the participant. The participant will be recorded using a Kinect or Zed 2i camera attached on top of the screen. Attached to the back of the screen is a Logitech webcam which provides the background image for the videocall. The videocall will take place using the AR based system. When looking at the screen, participant A sees participant B "projected" (3D point cloud image) onto the background image of the room participant A is in, so that he/she appears to be sitting opposite in the same room.

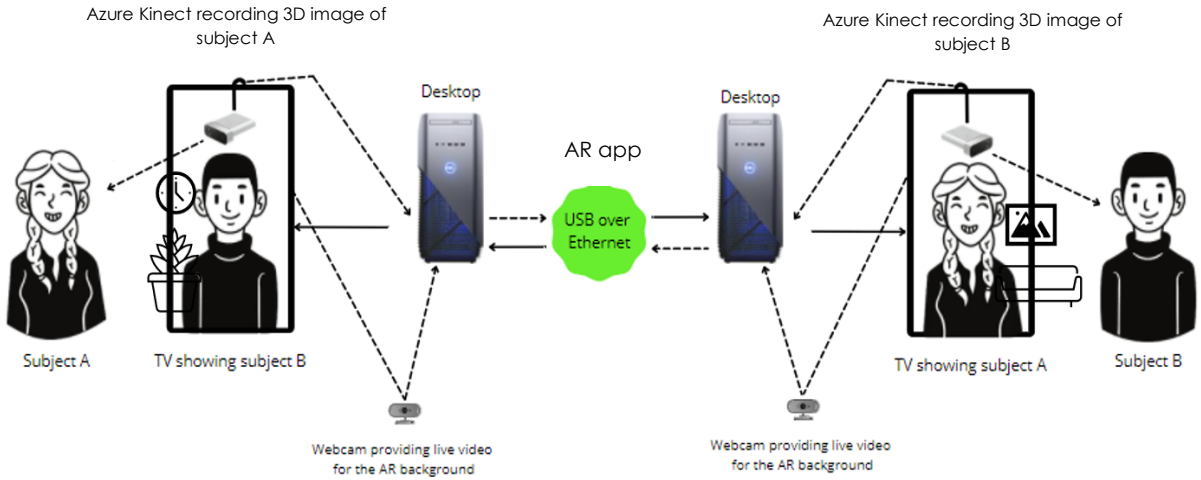


Figure 29. Technical setup condition II

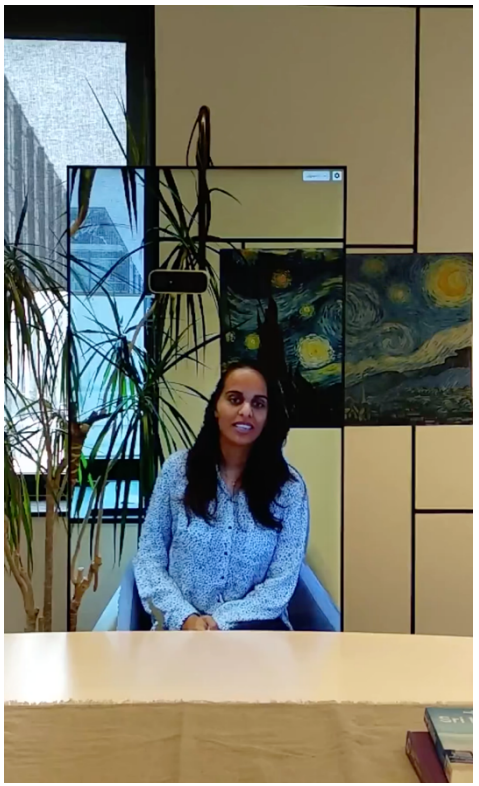


Figure 30. Video: appearance of AR-based tool

Analysis

IBM SPSS Statistics 28 (www.ibm.com) for MacOS was used to perform all statistical analyses. A significance level of 0.05 was used for all hypothesis testing. The normality of the data was assessed with Shapiro-Wilk tests. Between-group differences were tested using independent samples t-tests for normally distributed data and Mann-Whitney U tests for skewed data.

Thematic analysis was used to analyze qualitative data from open questions.

Results

Social presence:

HMSCQ: No significant differences were found in the general score of HMSC-Q between groups, neither in the copresence, intimacy, reasoning and behavioral scales. A significant difference (Z = -2.58, p = 0.01) was found in the credibility scale per condition, being the regular videocall tool rated higher than the AR tool.

Table 6. HMSCQ scores

	MS Teams communication	AR communication	*p* value
Total score HMSCQ	5.60 ± .81	5.50 ± .65	.60
Copresence scale	5.50 ± 1.14	5.00 ± .95	.76
Intimacy scale	5.50 ± 1.13	5.50 ± .99	.56
Credibility scale	5.50 ± .74	5.00 ± 1.32	.01
Reasoning scale	6.00 ± .69	6.00 ± .79	.79
Behavioral scale	5.75 ± 1.05	6.00 ± 1.13	.93

Note: Data are Median ± SD. Non parametrical tests were used due to

NMQ: No significant differences between conditions were found on the general scores of the Networked minds questionnaire, neither on the subscales.

Table 6. Networked Minds Questionnaire scores

	MS Teams communication	AR communication	*p* value
Total score NMQ	4.35 ± .48	4.37 ± .32	.87
Copresence scale	5.50[a] ± 1.14	5.00[a] ± .95	.29[b]
Attentional engagement scale	3.19 ± .44	3.17 ± .38	.89
Emotional contagion scale	5.50[a] ± .74	5.00[a] ± 1.32	.46[b]
Comprehension scale	6.05 ± .61	6.16 ± .55	.60
Behavioral interdependance scale	4.72 ± 1.35	4.91 ± 0.86	.60

Note: Data are Mean or Median[a] ± SD. [b]Non parametrical tests were used due

Differences on perceived social presence and perceived interpersonal closeness scores per type of relationship where explored, but this factor did not seem to have a significant effect.

Interpersonal closeness (IOS):

There was a median difference of one point on Interpersonal closeness score, higher using the regular videocall system (Md = 5.00). However, this difference is not statistically significant (Z = -1.36, p = .17).

Analysis of correlation between both social presence measures and interpersonal closeness (IOS) revealed a strong positive correlation (r = .72, p = <.01) only for the new social presence questionnaire (HMSCQ). There is no association between the Networked Minds score and perceived interpersonal closeness.

After this finding, coefficient of determination was calculated (R2 = .52). The variable social presence measured with the new HMSC questionnaire explains 52% of the variance of perceived interpersonal closeness measured with the IOS scale.

User experience (UEQ)

The total score of the user experience questionnaire did not show significant differences per system used, neither did its subscales.

Table 7. User experience questionnaire scores

	MS Teams communication	AR communication	*p* value
Total score UEQ	1.36[a] ± .66	1.30[a] ± .47	.61[b]
Attractiveness scale	1.83[a] ± ,68	1.83[a] ± .61	.67[b]
Stimulation scale	1.55 ± .81	1,25 ± .48	.18
Novelty scale	.51 ± 1.05	.82 ± .81	.34
Dependability scale	1.17 ± .74	0.97 ± .72	.40

Note: Data are Mean or Median[a] ± SD. [b]Non parametrical tests were used due

Nevertheless, according to the authors of this test, scores > 0.8 represent a positive evaluation, being scores above +2 or below -2 extremely unlikely. Therefore, both setups received a quite positive score. Also, these authors offer a benchmark to compare the scores of our product with results of other established products. On this comparison (table 7) we can see that both setups obtained Good and Above average scores on attractiveness and stimulation, the setup where the regular videocall tool was used scored below average on Novelty while the one using the AR tool scored Above average and the AR system scored below average on Dependability while the system using MS Teams scored above average. These results make perfect sense, being MS Teams a popular tool used for work-related meetings, and our AR tool a prototype, involving an emerging technology, that does not have a complete user interface yet.

Table 8. UEQ Benchmark comparison

	Scale	Mean	Comparisson to benchmark
Video-call system (MS Teams)	Attractiveness	1.83	Good
	Stimulation	1.55	Good
	Novelty	0.51	Below Average
	Dependability	1.17	Above Average
AR system	Attractiveness	1.83	Good
	Stimulation	1.25	Above Average
	Novelty	0.82	Above Average
	Dependability	0.97	Below Average

Quality:

The setup using the regular videocall system (MS Teams) was rated higher in overall quality and image quality.

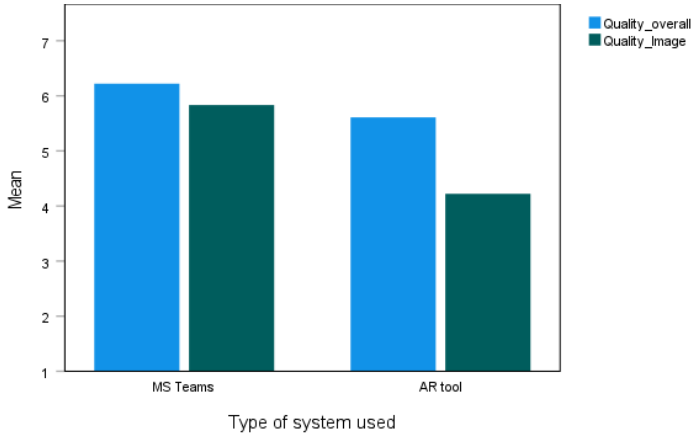


Figure 31. Perceived overall and image quality

Observation of behavioral patterns:

To enable a thorough analysis, behavior was coded into 7 factors:

1. Non-verbal communication: Hand gestures (count)
2. Natural postures and body movements (as in face-to-face): such as leaning head on the hands, crossing arms, playing with hands, etc. (count)
3. Interaction with objects (count)
4. Showing objects to the remote person (count)
5. Laugh (count)
6. Leaning back relaxed (duration)
7. Engagement: was the conversation interrupted with the end of the experiment? (Yes/No)

Figure 32. Participants using hand gestures (non-verbal communication)

Figure 33. Participants showing objects to remote person

Figure 34. Participants in a natural body posture (as in face to face)

The analysis of these factors (table 6) showed a significant difference in non-verbal communication, t (31) = 2.3, p = .03, (mean difference = 4.94). Participants used significatively more hand gestures to communicate using the AR tool (M = 11.88, SD = 5.63) than using MS Teams (M = 6.94, SD = 6.48).

There was also a significant difference (z = 2.33, p = .019, r = .40) on the number of times they laughed during the conversation, higher using the AR tool (Md = 7.00, n = 17) than using the regular videoconferencing tool (Md = 4.50, n = 16); as well as a marginally significant difference (z = 1.83, p = .07) on the time they spent leaning back relaxed, higher using the AR tool (Md = 210 sec, n = 17) than using MS Teams (Md = 12.5 sec, n = 16).

No significant difference was found for the other factors.

Table 9. Behavioral comparisson

	MS Teams communication	AR communication	*p* value
Hand gestures	6.94 ± 6.48	11.88 ± 5.63	.03
Natural body movement	11.88[a] ± 2.40	1.83[a] ± 2.26	.30[b]
Interaction with objects	4.38 ± 2.40	4.41 ± 2.26	.96
Showing objects	1.05[a] ± 1.35	0.00[a] ± 1.47	.12[b]
Laugh	4.50[a] ± 4.64	7.00[a] ± 3.69	.02[b]
Leaning back (time in sec)	12.50[a] ± 247.32	210.00[a] ± 372.99	.07[b]
Engagement	.14[c]		.71[b]

Note: Data are Mean or Median[a] ± SD. [b]Non parametrical tests were used due to skewed data. [C] Data is Chi squared value with Yate's correction

This being said, after reviewing participants' behavior while communicating in both conditions, it is fair to affirm that their interaction looks in general more natural and dynamic (as in face to face) than what we are used to see when using popular videoconferencing apps on a small size device (smartphone/tablet/laptop).

Qualitative results:

ITU score was supplemented with comments about pains, gains and suggestions, related to our AR-based communication tool, by answering three open questions:

What did you like about the system?

What would you change from the system?

Considering we are designing a system for domestic-personal communication, do you have any suggestions?

Most representative comments and keywords obtained:

Gains

Human-sized image	**Feeling connected**
Natural	**Face-to-face**
Less tiring	**Same room (background)**

“It wasn't as tiring as a video call through the computer, it felt more natural.” P22

“I liked also that the size of the person was "human-like", unlike videocall on a small laptop screen.” P24

“The feeling that you are in the same room.” P29

“It seems more that your partner is in front of you” P33

“Uses COTS (commercial off-the-shelf) consumer equipment” P35 (mentioned once but an important factor matching the insights from research on user’s needs).

Pains

No real eye-contact	**Background-Person integration (edges)**
Image distortions	**Inability to move (location in the room)**
Resolution	**Poorer image quality when moving**
Audio lag	

“The audio lagged, a better sync of audio and video would have made this experience much better.” P20

“It was strange that you could never really look each other in the eyes.” P21

“First basic thing: resolution, the quality of the image is noticeably poorer than standard videocalls.” P24

“If my partner was not leaning fully back in his chair, the image quality degraded” P26

“I think it does not work so well if you are not aligned in the center. You cannot move so much” P20

“Audio quality, the image lags, the image gets distorted when some movement happens, the background cropping is distracting.” P36

Suggestions/Comments

Image quality and size:
“The image edge fragmentation was a bit distracting… I think I might have preferred seeing my partners real background even if it did not blend with my environment.” P22

“I wonder whether the main point is just the "size" of the image. That's what makes the biggest impact for me. The attempt to simulate being in the same room is probably very effective if it's done perfectly, but when you start noticing the "cracks" you can't stop seeing it.” P24

One-to-one meetings:
“I'm not sure how this would work in a group setting, but for one-on-one calls I think this is a big improvement for communication as a deeper connection can be made.” P26

Better than current tools for a relaxed conversation:
“I would not prefer this above a physical meeting, but if it is more practical to have an online meeting, this is a great opportunity for a relaxed atmosphere to talk to each other. Better than a Teams or Zoom meeting.” P33

No need of 3D image:
“The quality of the person's edges. The 3D capture is unnecessary here; a 2D capture with background removal would suffice.” P32

Technological skills:
“I like the technology and I see how it could benefit a meeting if both parties have the same technological.” P36

Something worth mentioning from the AR condition is that, despite placing a moving item in the background, visible in and out the screen, 4 participants verbalized that they didn’t notice the background on the screen was real time, they thought it was a video previously recorded.

Discussion

- Participants did not consciously perceive a significant difference on social presence or interpersonal closeness between conditions, however, they used significantly more non-verbal communication, laugh more and spent more time leaning back (relaxed) on their seats while using TNO’s AR-based prototype to communicate. These are promising results for the application of AR on videoconferencing systems, that could be understood as a higher sensation of reality and naturalness, but they have to be carefully interpreted and supported with more research since, the most effective way of evaluating social presence should combine subjective and objective measures (Ijsselsteijn, 2001).

This increased use of non-verbal communication enriches the conversation by adding semantic, non-semantic and emotional information (Krauss et al., 1996), which can potentially lower "zoom fatigue", reducing cognitive load caused by constantly assessing attitudes, meaning and feelings.

- Some possible explanations to a more natural and dynamic conversation during both conditions of our experiment (in comparison with a regular videoconference on a typical small size device) are: Being hands-free allow them to gesticulate more, not having a self-view do not restrict their hands / arms movements (so much) to the camera's capture range and having a human-sized image with good-enough (image and audio) quality permits the person to lean back comfortably, while keeping the "eye-contact", not missing details.

- Both conditions actually received a medium-high social presence score in both scales, higher in the new HMSCQ (>5). As previous work suggests (Ahn et al., 2014; Ishida et al. 2011; Koh, 2010) this positive score can be caused by the human-sized image of the remote person, presented in both conditions and would be lower when using a smaller size display. This, supported as well with our qualitative results: "I liked also that the size of the person was "human-like", unlike videocall on a small laptop screen." (P24); leads to think that display size has a major impact on the perception of social presence. Further research is currently being performed at TNO to confirm this specific hypothesis.

- The finding of a strong correlation between one of the measures of social presence (HMSCQ) and interpersonal closeness. Leads to think that the old Networked Minds questionnaire is no longer sensitive enough when measuring aspects of social presence using social XR technology, and that there is indeed a strong correlation between these two factors, with implications on emotional connection of people in a long-distance relationship, but further research is needed.

- When it comes to overall User experience evaluation, the UEQ (Schrepp, 2015) showed no differences between conditions, probably due to the similarity of both systems in terms of appearance (hardware, size, etc.) and the lack of interaction with a user interface (the call was initiated in both cases by the experimenter and participants only had to perform a conversational task). However, UEQ scores were very positive in both scenarios, supporting therefore the adoption of UX "gains" identified in this experiment as inspiration for the design of future VC systems. The highest score (M = 1.83) on this questionnaire correspond to the Attractiveness scale, an overall impression of the product, aligning with the score obtained on intention to use: 5.44 / 7.00

- The lower (overall and image) quality score obtained by our AR-based prototype can be explained by some visual artifacts happening on the image of the remote person, and Microsoft Teams being a well-known commercialized tool which quality has been improved over time. Real time 3D (point cloud) image capturing/rendering was the main cause of the visual artifacts and poor resolution when moving and, since the issues caused by this technology overweight its

benefits on a flat screen, it would be recommended to avoid it until it can provide a comparable image quality.

- Lastly, insights extracted from qualitative research point us the direction to follow in order to improve the user experience and feeling of social connectedness of new VC systems for the domestic environment, where extended reality has a big potential, but not in isolation. The use of emergent technologies must be aligned with improvements on image and sound quality, displays/devices (size and mobility), eye contact, natural interaction, features for a fluid and enjoyable social connection, and an intuitive cross-generational user interface.

Conclusion

In this study we compared the user experience of an AR-based video-communication tool with a regular VC tool, for social interaction. No differences were found on subjective measures (questionnaires) of overall UX, and more specifically perceived social presence, among both communication modalities. On the other hand, behavioral analysis showed an increased amount of hand gestures (non-verbal communication), as well as a more relaxed attitude, laughing and leaning back on their chairs more often) when using our AR-based tool. Therefore, we can just partially conclude that AR provides with a better experience than current VC tools, given the same conditions, at least in terms of naturalness. This is supported by a high intention to use expressed by participants after experiencing the AR-based tool.

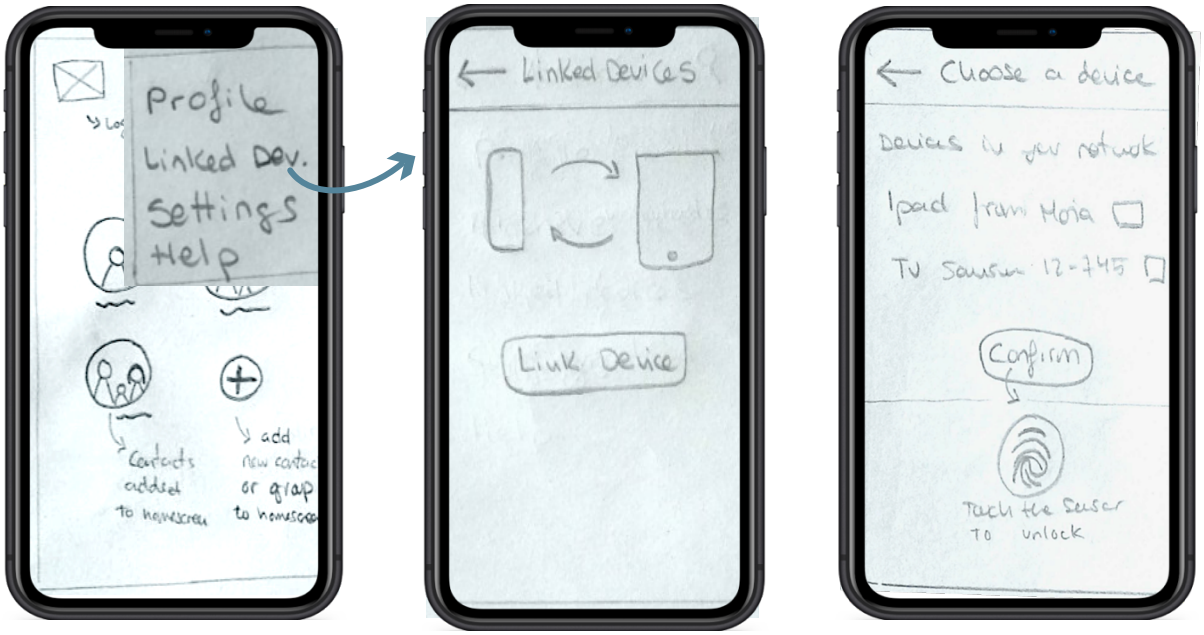
Both scenarios actually scored very positively on social presence and general UX, as well as, promoting a more dynamic and natural-looking conversation than what we can observe during common VC on small size devices. This leads to consider that factors present in our experiment, different from the AR implementation itself, are having an influence. Therefore, to achieve a more advantageous holistic UX for video-communication systems (for domestic use) and compete with current tools, the implementation of extended reality must be integrated in a good strategy involving other functional, technical and contextual factors.

APP SKETCHES COMPLETE

Onboarding

Incoming videocall

Menu: Linked devices



INFORMED CONSENT USABILITY EVALUATION

The purpose of this study is to evaluate the interface of a new video communication tool. Your participation will help us find areas for improvement in it.

- Your participation in this study is voluntary
- You can take a break whenever you want, just tell the researcher.
- You can leave the study whenever you want if you think it is necessary.

Information we want to collect

We are going to ask you to perform various tasks with a prototype of the app. You will be watched as you perform these tasks and the session will be recorded. We would like to ask you to think out loud while you are performing the tasks.

How we ensure your privacy

The recordings will also be viewed by the tutor of the subject, but the content will be treated as confidential material and under no circumstances will it be shared outside the scope of this evaluation. Your comments can be recorded in the final report of this test, but will be completely anonymous. That is, your personal data will not be named at any time in the report of this investigation.

If you wish to withdraw your consent, please contact the person mentioned below and they will destroy your personal data and any material we have from the session. Otherwise, your data will be deleted after the completion of this project.

Marina Alvarez

maralvarezmer@uoc.edu

0611410728

I consent to:

(Please check the boxes that apply)

☒ My session is recorded

☒ The recording is used for academic purposes

Name: Beatriz Ramos

Place: Eindhoven

Date: 17-05-2022

Signature