# Social Augmented Reality Application

## Enhancing Remote Collaboration in
## Physical Work Context

Sanna Reponen
Master's Thesis

MA in New Media Design and Production
Department of Media

Aalto University
School of Arts, Design and Architecture
2016

**Aalto University: School of Arts, Design and Architecture**

**Author** Sanna Reponen

**Title of thesis** Social Augmented Reality Application: Enhancing Remote Collaboration in Physical Work Context

**Department** Department of Media

**Degree programme** MA in New Media Design and Production

**Year** 2016 **Number of pages** 99 **Language** English

**Abstract**

Effective, collaborative problem solving in physical work requires building common ground and active awareness of the situation. The main hypothesis of this thesis is that mobile video conversations augmented with an on-screen drawing feature are beneficial for problem solving and communication in physical work contexts.

The thesis consists of the description and analysis of the design process of a mobile video application called Social Augmented Reality (SoAR) and the results of related background and user research done in the construction, facility maintenance and quality management sectors. Mapping how the users experience solving physical tasks with SoAR and especially how the drawing feature is used are the key goals of the study. Methodologically the thesis is a combination of qualitative field and interview studies and digital tool design.

The iterative design process of SoAR is founded on the research-based design framework developed by Teemu Leinonen (2010). The design and development steps are described and usability research is conducted in a controlled environment in order to learn how spoken language and drawing on screen intertwine when collaborating remotely using SoAR. Usability research indicates that pointing by drawing is experienced as effective and that drawing on screen allows compact use of spoken language.

Data on current communication practices in construction, facility maintenance and quality management is collected through interviews and job shadowing. Augmented video calls with SoAR are piloted in actual work contexts, and the field tests are analyzed on the basis of recorded SoAR calls and final interviews with the users. Augmented video calls appear to have a lot of potential in enhancing remote collaboration due to effective pointing at task objects and locations.

SoAR is an open-source mobile application and one of the products developed in the Learning Environments research group of Aalto University, Department of Media, as a part of the Learning Layers EU project. The original concept of the application was developed by Doctoral Candidate Jana Pejoska.

**Keywords** research-based design, augmented reality, remote collaboration, CSCW, construction, facility maintenance, quality management

CONTENTS

## ACKNOWLEDGEMENTS

This thesis would not exist without the Learning Environments research group of Media Lab, Aalto University. I am grateful to Teemu Leinonen for the research opportunity and the resources I have been able to utilize. I owe a special thank you to Jana Pejoska for placing her trust on me: I am happy I got to develop further her concept of Social Augmented Reality, to the best of my ability. I would like to thank my instructor Marjo Virnes for her presence, collaboration and precious advice, and Matti Jokitulppo for being my partner in crime - the actualizer of SoAR. I am also grateful to Jukka Purma, Eva Durall, Tarmo Toikkanen, Merja Bauters, Samuli Raivio, Lassi Veikkonen, Leo Nikkilä and Gerhard Molin from LeGroup and my supervisor Rasmus Vuori from Media Lab.

A big, loud thank you to all Media Lab test users and Mlab people in general. I am especially grateful to Laura Uusitalo for collaboration and inspiration in a number of occasions along the way. Thanks paper prototypers Riku Erkkilä, Yun Hsuan Huang, Soujanyaa Boruah and Andrei Duvan Rodriguez.

There are also a number of people from the Learning Layers project and the industry whom I owe a debt of gratitude: Melanie Campbell and the apprentices from BauABC, Markus Manhart and Stefan Thalmann from the University of Innsbruck, Pekka Kämäräinen from ITB, University of Bremen, Heikki Qvick from Skanska, Jussi Kymäläinen, Mira Bartholdi, Sami Immonen, Juho Laurila and Tomi Hardén. Thank you Tapio Uusikartano from Nokia for mentoring me during the final year of my studies and the thesis project.

Thank you Tero Toivola for supporting me no matter what.

## LIST OF FIGURES



## LIST OF TABLES

# 1. INTRODUCTION

Augmented reality (AR), the digital information overlay on the real world, is considered to be the next big thing in digital technology along with more immersive virtual reality (VR) applications, according to finance and technology corporations like Deloitte, Goldman Sachs and Gartner. The economic and practical impact of AR and VR is expected to be similar or even bigger than mobile technologies and especially smartphones had in 2010's. The biggest hype revolves around emerging hardware platforms, specifically head-mounted displays (HMDs), which are anticipated to mature and enter the market with full force in a few years' time. (Kunkel et al. 2016, Uzialko 2016, Bellini et al. 2016.)

Future AR use cases in for example entertainment, education and retail are abundant, but even if AR is also seen to have potential to revolutionize work life in general, descriptions of how exactly this is will happen often remain vague. According to Kunkel et al. (2016): "Deploying augmented interfaces that pair with connected devices, sensing objects, and relational data can deliver task-specific information to workers in the field in context and on demand." Remote conferencing will be improved, because AR and VR have "the potential to allow a global workforce of remote teams to work together and tackle an organization's business challenges. No matter where they are physically located, an employee can strap on their headset and noise-canceling headphones and enter a collaborative, immersive virtual environment."(Sena 2016.)

What is required in my opinion is sufficient understanding of the actual work contexts in which AR is supposed to land in the next few years. Enhancing remote collaboration is a key interest in this thesis, and one of the first observations I made in the course of the work was, that mobile devices and especially video call applications would already provide users with a good basis for remote collaboration over a shared view. However, even if the benefits of mobile video tools were acknowledged in the organizations that participated the study, none of the test users from different lines of physical work had thought of making video calls outside the conference call context.

Thus, I will turn the attention to current mobile video communication and mobile AR and learn as much as possible about the advantages and challenges that such collaborative technologies pose for real-life work. Studying the framework of physical work and how employees collaborate with and without technology will pave the way for better design and adoption of AR hardware, as well. According to Gartner, by 2019 20 % of mid-sized organizations with field services will be involved projects that utilize HMDs (Robinson et al. 2016). Considering such a slow adoption expectation of AR hardware, it is worthwhile to make use of what already is out there – preferably in a simple, adoptable and user-centric fashion. The aim of this thesis is developing a mobile, augmented video call application that is validated with user research through and through. The application needs to be simple enough to be used flexibly in as many physical work contexts as possible. The data collected from the user tests of the application will provide a set of guidelines for future, collaborative augmented applications that rely on shared view.

## 1.1 Background

For this thesis I studied how mobile video conversations augmented with drawing and pointing features helped problem solving and communicating in work context, more specifically in construction, facility maintenance and quality management work.

The thesis is a combination of qualitative field studies and digital tool design: I make use of descriptive, discovery-driven activities as well as generative, design-driven activities (Wobbrock, Kientz 2016). The prototype developed and used in the thesis project is called Social Augmented Reality application (SoAR). I have been responsible for the design choices and the user research of the prototype.

SoAR is an open-source mobile application and one of the products of the EU project Learning Layers. The project has taken place in 2013-2016 and one of the partners has been the Learning Environments research group (LeGroup) of Aalto University, Department of Media. Doctoral candidate Jana Pejoska developed the concept of SoAR in 2014, and I worked with the prototype in Jan-Aug 2016. In its current state the app is a working Android prototype, developed by research assistant Matti Jokitulppo.

In the thesis I describe and analyze the design process of SoAR, which is founded on the research-based design framework developed by Teemu Leinonen (2010). In order to find out how well SoAR is suited for professional use, the prototype has been tested in collaboration with construction, facility maintenance and quality management workers. Data has been collected through interviews, job shadowing and on-site testing of the prototype. Getting to know the users and their work has helped me find out what the current communication practices in the workplace are like and which specific situations could be handled with SoAR. I have conducted usability tests to understand how drawing improves remote collaboration and to improve the design iteratively before on-site user testing. The design and research activities conducted during the thesis project are presented in Figure 1.

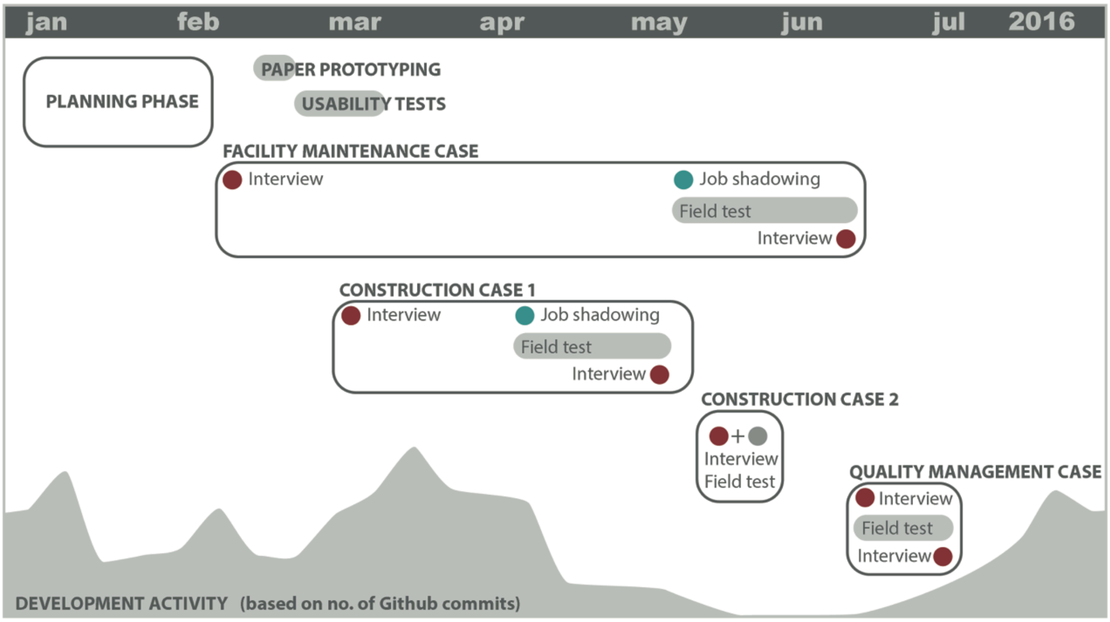



**Figure 1:** Design and research activities conducted during the thesis project

Future research on the topic should aim to find out whether tools like SoAR have the potential to change collaboration and help-seeking practices in the workplace and how emerging AR hardware and new mobile sensors like depth-sensors and 360° cameras will transform work-related visual communication.

## 1.2 Research Objectives

I set out with the main hypothesis that mobile video conversations augmented with an on-screen drawing feature are useful for problem solving and communication in physical work context.

I aimed to find out what the current communication practices in construction, facility maintenance and quality management sectors are like and how augmented video calls could change these practices. Mapping how the users experience solving physical tasks with SoAR and especially how the drawing feature is used were key goals of this thesis. In addition, I described the design process of SoAR and assessed how well it succeeded.

## 1.3 The Structure of the Thesis

In *Chapter 2* I present related research and projects on augmented reality and describe how it has been used especially in construction and maintenance. The importance of shared view in remote collaboration is discussed in relation to earlier research in the field of computer-supported cooperative work (CSCW). Earlier studies on augmenting video feed with drawings in collaborative tasks are introduced.

*Chapter 3* introduces the research-based design method and the timeline of the design process of SoAR, including both past activities and those completed during this thesis project. The data achieved from various research and design activities is described here.

*Chapter 4* presents the analysis of the usability research conducted in a controlled environment. The usability test participants accomplished a set of tasks in remote collaboration with SoAR. The discussions during the tasks were reviewed using conversation analysis methods in order to understand how drawing on screen and verbal communication support each other in video-based collaboration.

The topic of *Chapter 5* is testing SoAR in actual, physical work context. Interviews, observations and field test documentation were analyzed thematically to assess whether SoAR has the potential to improve remote collaboration in construction, facility maintenance and quality management sectors and in physical work in general.

*Chapter 6* concludes the thesis with general design outcomes and an assessment of how well the research objectives of the thesis were fulfilled.

# 2. RESEARCH CONTEXT AND RELATED WORK

## 2.1 Augmented Reality in Construction and Maintenance

Augmented reality (AR) is defined here as a technology that allows the simultaneous viewing of real world, either direct or mediated, and virtual objects that are superimposed or otherwise embedded to the view (Azuma 1997). AR also has the potential to augment for example hearing and sense of touch, even if most applications aim to enhance vision.

In maintenance and construction work, it is essential to be able to compare the physical work surroundings and the objects located in it to instructions and visual models, e.g. structural charts, connection diagrams and 3D models. Typically this data is viewed on a display or paper. AR applications afford simultaneous viewing of the physical context and the models as overlay, as well as tracking changes. Baird and Barfield (1999) have demonstrated that using AR in assembly can significantly reduce the amount of errors and speed up completion times. Tang et al. (2013) have come to similar conclusions, and note that AR instructions also diminish the mental workload in assembly work.

The first work-related AR concepts from the early 1990s were aimed specifically at enhancing maintenance and manufacturing work (Caudell, Mizell 1992, Feiner et al. 1993), and concepts and prototypes for construction started to appear a few years later (Klinker et al. 2001, Schall et al. 2009, Woodward et al. 2010). Maintenance and manufacturing have apparently been the first fields to explore because in many cases the work flow can be broken into relatively clear subtasks and processes, whereas construction sites can be challenging test areas due to constant change in the work environment and harsh weather conditions.

Apart from novel human-computer interaction, AR also affords new ways to collaborate. The first collaborative AR systems in which the users were able to work on digital content in the same physical environment were prototyped in the 1990s (Szalavári et al. 1998, Billinghurst et al. 1998, Rekimoto 1996). AR is noted to reduce separation between the actual tasks and the communication related to them, as AR has the capability to enhance communication cues typical to face-to-face situations (Lukosch et al. 2015a). Remote collaboration can become more immersive and functional when AR is exploited to enhance the experiences of sharing, exploring, referencing and manipulating the physical task context (Gauglitz et al. 2014a, Gauglitz et al. 2014b).

Caudell and Mizell (1992), in their report of the head-mounted display prototype they developed for assembly work at Boeing, note that the graphical elements in an immersive virtual reality (VR) system can be considerably complex compared to an AR system, the latter of which can thus be driven with less computing power. AR is an obvious choice instead of VR for most maintenance and construction tasks due to occupational safety, and relatively low power consumption of the systems supports developing mobile solutions, which is essential as the workers need to constantly move from one place to another.

Recently, MARIN and MARIN2 projects of University of Turku have concentrated on developing intuitive visualization of structural plans like CAD models for marine industry and construction. Communication between the planners and the workers at the construction site, as well as keeping the plans up-to-date, could also be enhanced by connecting technical data to the actual site with AR. (Lehtonen 2016, Helle et al. 2014.)

Apart from connecting data to a specific location only, Zhu et al. (2013) call for AR applications that are able to build awareness of many other contextual clues including data collected of the users' preferences and their level of expertise. He also notes the need to increase the level of user authorship in maintenance AR, which would benefit the process of building more context-aware systems. In addition, context-awareness and authoring done by users would help tackle the challenge of the changing environment in construction.

The work contexts studied in this thesis, construction and facility maintenance, require the workers to move from place to place. Because of this, non-mobile AR applications were out of the focus of this thesis. With non-mobile AR I mean applications that do not require the user to wear or carry any technological devices but where the technology is built into the surroundings. Also desktop computers can be considered as non-mobile devices with AR capabilities.

The variety of different AR capable mobile devices is on the rise. Along with touch-screen based mobile phones, tablets and smart watches, a number of see-through head-mounted displays (HMDs) by for example Microsoft, Epson and Vuzix have been announced lately (Härkänen et al. 2015). HMDs allow the users to interact with gestures, operate hands-free and, at best, view the real world and the virtual components seamlessly.

Among the advantages of using handheld devices for AR applications instead of HMDs are availability, minimal intrusion and social acceptance (Zhou et al. 2008). Availability and thus the possibility to immediately utilize the Social Augmented Reality application (SoAR) in real work environments were the key reasons for sticking to touch-based mobile devices within this study. It was also possible to conduct user tests in natural work situations, not ones that are staged around a technology that would otherwise not be present at the work place. For example, modern HMDs are not yet suitable for construction work due to the requirement to wear safety glasses and hearing protection equipment at the site.

## 2.2 Remote Collaboration with Augmented Video Applications

### *2.2.1 Collaboration over a Shared View*

Advances in information technology and the development toward more and more global economies have made geographical distribution of work common in numerous industries. Remote collaboration is dependent on Internet-based communication methods like email, instant messaging, file and screen sharing services. As bandwidth and computational power in both computers and mobile devices have increased, live video communication with affordable tools like Skype has become commonplace both in work context and in everyday life (Wikipedia 2016).

In construction and facility maintenance sectors, collaboration among a network of subcontractors and stakeholders who are not constantly present at the physical work locations takes place on a regular basis. The collaborative work situations in this thesis typically require making a visual assessment of the circumstances and sharing it effectively with members of the cooperative network, and augmented video conversations are studied as a means to improve communication. Visual information helps the conversation partners reach and maintain a *common ground*, which a prerequisite for *situation awareness* and actions that follow on the basis of the current situation. With common ground I mean the basis of collaboration, the mutual understanding that consists of a set of beliefs, assumptions, bits of knowledge etc., which is built in the process of *grounding* where the people who are collaborating build and keep up such an understanding (Clark, Wilkes-Gibbs 1986, Baker et al. 1999). Situation awareness means a person's state of understanding of the dynamic elements related to the task at hand in a specific environment – good situation awareness results in decisions that are likely to be appropriate (Endsley 1995).

Both free time and work-related video communication take usually place in facial view formation. However, studies suggest showing the surroundings and the objects located in it is more useful for effective cooperation than concentrating only on faces: facial view is not optimal for building common ground and situation awareness among the conversation partners (Licoppe, Morel 2012, Gergle et al. 2013).

When comparing audio communication to video communication that only shows the faces of the collaborators, without other visual cues, there appear to be no significant benefits to video – on the contrary, even if users often prefer seeing their conversation partner, video connection may result in more overlapping utterances and breaks in the conversation than audio-only connection (Sellen 1995, O'Malley et al. 1996, Anderson et al. 1997). Apart from being able to see the physical task context over video, it is also essential to view the partner's actions (Daly-Jones et al. 1998). It seems evident that seeing the physical surroundings is more essential than seeing the conversation partner in immediate work situations related to construction and facility maintenance.

In a controlled user test, Gergle et al. (2013) have compared pairs who collaborate on a puzzle task over shared-view video connection or only relying on audio communication. The performance was consistently and considerably faster when the collaborators had a shared view. Also, the shared view was even more beneficial the more linguistically complex the task was. Earlier studies back up these findings (Clark, Krych 2004, Gergle et al. 2004, Fussell et al. 2000, Kraut et al. 2002). Despite their usefulness, it appears video conversations cannot easily replace collaboration in the same physical space: working side-by-side appears to be preferred by workers and more effective than distributed work (Fussell et al. 2000, Biehl et al. 2015).

All in all, collaboration over a video connection is a viable solution in remote collaboration as long as the partners can share the view of the task at hand and of each other's actions, but it has limitations compared to working side-by-side. These limitations could be mitigated significantly, should a video communication tool support referring to task objects by effective pointing possibilities and high quality visual cues.

In work situations that require mobility, it has been suggested the collaborators will benefit if the remote, the person located away from the situation can somehow control the view that is streamed from the site. Such control is taken to improve the sensation of "being there", which is also referred to as telepresence. (Rae et al. 2014, Gauglitz et al. 2014a.) Telepresence can be enhanced by mapping the visible work area with a camera and then allowing the remote person control the view within the limits of the mapped area (Gauglitz et al. 2014a). Also, the remote user may suggest a new viewpoint to the conversation partner by moving her mobile device - the gyroscopic data of the remote's mobile device is visualized on the other user's screen as a hint where to turn the camera (Jo, Hwang 2013). Another possibility is robot-mediated communication, where the remote has control over a robot camera: the robot can be steered freely in the task location to gather more visual cues and to reassert the telepresence of the remote (Rae et al. 2014).

According to Gergle et al. (2013, p. 7), creating communication technologies to mediate visual information involves a trade-off of features – it is not possible to utilize all technological features that would potentially bridge the flaws of "not really being there" simultaneously. Also, the necessity of different features depends on the tasks at hand. Technological choices will have an effect on frame rates, voice synchronization, alignments of the views and perspectives and the control of these, to mention a few. Allowing the remote collaborator to control the view is an asset in many situations, but mapping views will have a cost on the streaming quality as the amount of visual data and calculations increases. Bringing in remotely controllable cameras or robots is currently not a practical solution in real-life work situations in construction and facility maintenance: the workers move from place to place, outdoors and on rough terrain. The solutions need to be portable, durable and off-the-shelf. Because the work situations studied in this thesis are largely unpredictable and often require quick decision-making, pointing by drawing is considered as the key augmentation for reaching the optimal common ground and situation awareness, as fast as possible – enhancing the telepresence or view control of the remote, useful as it may be in many contexts, is thus not a priority.

#### *2.2.2 Augmenting Video Stream with Pointing and Drawing*

When working on a physical task in co-located settings, collaborators tend to use gestures to clarify the verbal communication they are engaged in. Gestures can generally be divided into *pointing* and *representational* gestures. Pointing is used to refer to objects and locations, whereas representational gestures are used to describe forms of objects and actions that need to be done (McNeill 1992, Bekker et al. 1995, Fussell et al. 2004). According to Fussell et al. (2004), full ability to use pointing and representational gestures is among the key reasons why co-located collaboration is more effective than video-mediated collaboration. When collaborating over video stream, users tend to adjust cameras and their behavior in order to enable pointing, which takes time and shifts the focus of the collaborators towards the adjustment maneuvers instead of the task at hand. Also, shared video stream view is always more limited than the view of co-located collaborators, which makes it all the more important to use gestures like pointing to confirm the remote collaborators have established common ground (Huang, Alem 2013).

One approach to supporting gestural information in remote collaboration is to project or simulate actual hand movements in the collaborators' view (Huang, Alem 2013, Robert et al. 2013). However, as SoAR relies on touch screen devices, pointing and drawing directly onto the video screen were chosen as the method to allow gesturing between the collaborators. Drawing on video has been recognized as a way to mitigate physical distance in distributed work quite early on. In the beginning of the 1990s, Xerox Palo Alto Research Center created prototypes that utilized video connection and shared drawings for distributed design work. The prototypes conveyed collaborative drawings and hand gestures along with audio conversations (Minneman, Bly 1991, Tang, Minneman 1991). These prototypes along with later work by e.g. Kato and Billinghurst (1999) are reminiscent of working on a shared virtual whiteboard, where the focus is on the drawings themselves, not the working environment.

Bauer et al. (1999) developed a wearable videoconferencing system to test whether a feature they call *reality-augmenting telepointer* improves the performance of pairs working remotely on a network maintenance scenario. The remote worked on a desktop computer and controlled a telepointer in the shared view to give instructions, while the local user provided the shared video view with a head-mounted camera-display unit. The system had an option to freeze the image to compensate the shaking HMD view. The study showed a strong preference for using pointing for remote guidance. Pointing reduced the amount of verbal instructions when compared to performances based only on audio and video. Later on, several studies have confirmed users perform faster and more reliably when they collaborate over annotation-augmented video instead of normal video stream (Kim et al. 2013, Gauglitz et al. 2014b, Fussell et al. 2004). In all of these studies, the remote worker or helper used desktop technology, whereas the so-called local user, who was performing the actual tasks, had a mobile setup. Even if the local user had the possibility to move thanks to the mobile gear, the tasks in these studies were stationary and did not require moving from one place to another. Drawing or pointing activity was limited to the remote worker's interface.

Fussell et al. (2004) tested two different video-mediated pointing methods on a toy robot assembly task: pointing with dots and drawing free-hand. According to the results, pointing with dots was too limited to improve task performance in comparison to audio-and-video-only communication. Free-hand drawings, on the other hand, allowed a wide range of gestures – Fussell et al. claimed the performance and communication is practically equal to collaborating in the same space. Free-hand drawings were also found superior to simple cursors or dots by Kim et al. (2013). However, since the tasks used in these studies were static assembly tasks with construction blocks, it is not evident how well the results relate to real work contexts.

Jo and Hwang (2013) introduced Chili, a mobile draw-over-video system and suggested the users of a system like this should be provided with the same user interface, regardless of what the role in the research task is. In a two-user setting, the drawings should be bidirectional and both users should carry mobile devices. Like Jo and Hwang, I assume a remote user may become a help-seeker at any given point, as she might need to help a colleague and suddenly need help or advice from yet another worker. The same application with identical features and hardware are expected to work for the all SoAR users.

All in all, previous research points out there are clear benefits to implementing a drawing feature in a video collaboration application. The benefits of a fully mobile system have not been studied extensively as, probably due to the need to control the test settings as well as possible, test tasks have been often been carried out in a stationary setting like Lego or model assembly.

Also, there is relatively little research on how such applications could be used in real work contexts and how useful different drawing gestures are in sudden, unprepared situations. Lukosch et al. (2015b) have conducted a study on an HMD-based, shared view AR system for enhancing the work of security personnel and forensic investigators. Domova et al. (2014) studied the use of a mobile video interface augmented with drawing capabilities in a wastewater treatment plant. The results from both of these studies suggest that drawing and annotation gestures on video stream improve collaboration with a remote expert. However, also in these studies the role of the remote was expected to be fixed, as the remote only worked on a desktop interface different from the one used in the field.

## 2.3 Other Applications with Augmented Live Stream

In addition to research projects, mobile applications with drawing and pointing augmentations have been developed and released also in the private sector.

A common use for commercial augmented video applications is free time socializing. Typically such applications include minigames to be played with the conversation partner and multiple color choices for drawing on screen. The initial setup is face-to-face view. Instant messaging is often included in the apps. JusTalk and Rounds are fairly popular examples of video chat apps that include a drawing feature.

Augmented video streaming applications that aim to enhance remote collaboration are not nearly as popular as freetime apps. Even if the online application store Google Play is not the only place to download such apps, the download numbers are indicative: Rounds has been downloaded 50 – 100 million times, whereas work-oriented ones like POINTR by Delta Cygni Labs and Onsight Connect by Librestream Technologies have the download count of around 1000. The current version of POINTR makes mostly use of premade symbols, called pointer annotations, and according to the developers, special attention has been paid to usability despite network disruptions and slow speed connections. Onsight Connect is a collaboration platform aimed at corporate use, especially remote collaboration in field work.

There are also several call applications that enable drawing on a still view instead of video stream. Sessio by Flashwalk provides live audio and drawing takes place over a view shared by the conversation partners. Skype by Microsoft and Hangouts by Google can also be augmented with collaborative whiteboard tools, which allow streaming the audio and drawings in a similar fashion to Sessio.

# 3. SOCIAL AUGMENTED REALITY: DESIGN AND RESEARCH PROCESS

## 3.1 Research-Based Design

The design process of SoAR follows the research-based design method as outlined by Teemu Leinonen (2010). This form of participatory design takes advantage of research as a means to find out how to reach the most viable outcome – the best possible artifact or tool. The aim of the design process is, however, not only the artifact but also an effect on reality: a perceived change in the working culture, the community of people, the everyday life. Such a change may sometimes be out of reach, but the purpose of reaching it, in effect, requires the design process to be iterative. The objective of SoAR – as an artifact and as a design process – is to lower the threshold to seek and give guidance and, as a result, enhance the culture of communication and mutual assistance at workplace. Thus, at best it has the capability to support workplace learning and professional development. SoAR may mitigate the effect of distance and different backgrounds and vocabularies of the workers, as it provides a visual means to communicate and emphasize details. It is also possible that SoAR has a positive effect on workplace safety as cooperation in acute situations is enhanced.

The phases of the iterative process are *contextual inquiry, participatory design, product design* and *prototype as hypothesis* (Figure 2).

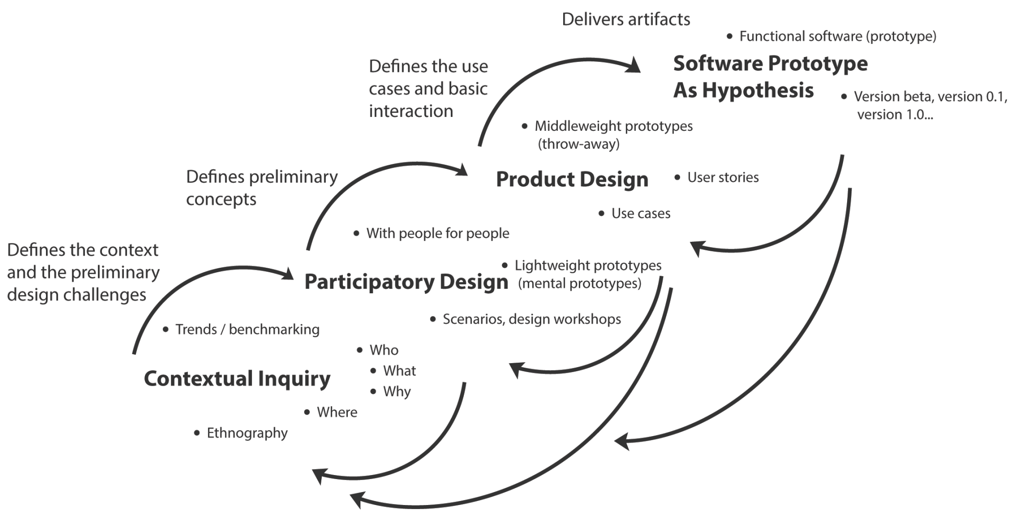



**Figure 2:** Research-based design method (Leinonen 2010)

Even if the phases of the design process can be described as linear, they are in reality overlapping, and it may often be necessary to track back to previous stages. My work was mostly concentrated on the product design and prototype as hypothesis phases, but I draw material and influence from the other phases as well – both in the form of results of earlier research on SoAR but also as methods I found necessary to do my own research, to validate my design choices.

Contextual inquiry takes place first in the process. It is required for understanding who the design is meant for and what is supposed to be achieved with it, what the socio-cultural framework of the design is and what trends the design is related to. Even if some contextual research on SoAR had been done already in 2014, I needed to go back to this phase to understand the specific work and communication environment my test users were involved in.

Participatory design as a general term covers a variety of approaches, but it is here defined as the stage where most input is collected from stakeholders. The methods to gain input are engaging – from co-design workshops to co-created scenarios and low-tech prototypes. The participatory design phase was partly overlapping with contextual inquiry in my study: I gathered as much user input as possible to ground each design step, but the main concept of the augmented video call application remained the same. No new prototype ideas or scenarios were created at this point of the project.

In the product design phase some distance is taken to the stakeholders and the results of the previous phases are turned into interaction models and information architecture. Use cases and rough prototypes are created. The most distinct design phases in the SoAR design process occurred before each user test cycle: before usability testing and again before field-testing.

Prototype as hypothesis is the phase where functional prototypes can be released and tested in the intended environment to find out whether they have the potential to solve the challenges the design process seeks to address. The field tests conducted in the work places were the key phase where the validity of the SoAR prototype as a solution for real-world issues has been put to the test.

## 3.2 Early Development: Towards the Proof-Of-Concept

SoAR was developed as a part of the Learning Layers project (EU, FP7, 2012-2016), which had the objective to develop mobile tools for informal learning in the workplace (Learning Layers 2016). Learning Environments research group (LeGroup) was responsible for the Learning Layers work package 4, which consists of digital prototype development and, in early stages of the project, providing design expertise to partners in the form of co-design workshops and guidelines.

The first steps towards an AR-based communication tool suitable especially for construction work were taken in Sep - Dec 2014. The contextual inquiry and participatory design phases occurred side by side, as benchmarking and mapping the AR needs in construction took place in co-design workshops with project members and stakeholders, and as participatory research was done at three construction sites in Finland.

Situations that require collaboration, visual assessment and agreement between a remote and physically present worker were identified as key issues that could be addressed with an AR solution. Communication gaps were noted - these were often related to new workers and how they reach a sufficient level of knowledge to be able to function at the site. Generation gap was a cause of worry, as experienced workers have a lot of knowledge to pass on and, respectively, younger generations have adopted tools and technologies that have not been considered easily accessible to the older workers. According to the representatives from construction education, the communication between vocational schools and construction sites with apprentices should be improved. Specific conditions for AR design for construction are constant change in the physical environment and harsh weather: both the content and the tools used should adapt to the conditions. The most viable AR design would thus address the need to build professional identity and knowledge pool, and on the other hand, enhance communication and help seeking in situations that require immediate attention. (Bauters et al. 2014, Pejoska et al. 2016.)

In the product design phase, the concept of SoAR took form to consist of a social network component and tools for communication (Figure 3). In order to strengthen professional identity, SoAR could be used to accumulate information on a worker's work history and experience. Profiles are also beneficial for finding the right contact when seeking help. The communication features would consist of instant messaging and vision sharing, i.e. augmented video conversation, the latter of which is considered the main feature of the application. The asset of instant messaging is the possibility to share data like images, locations and text. (Pejoska et al. 2016, p. 479.) The application was designed not to include premade graphical content in the way AR applications typically do, as the constantly changing physical environment at a construction site poses a challenge for keeping content-heavy AR up-to-date. All the content is to be created during the video conversations or as a result of activity in the social network.

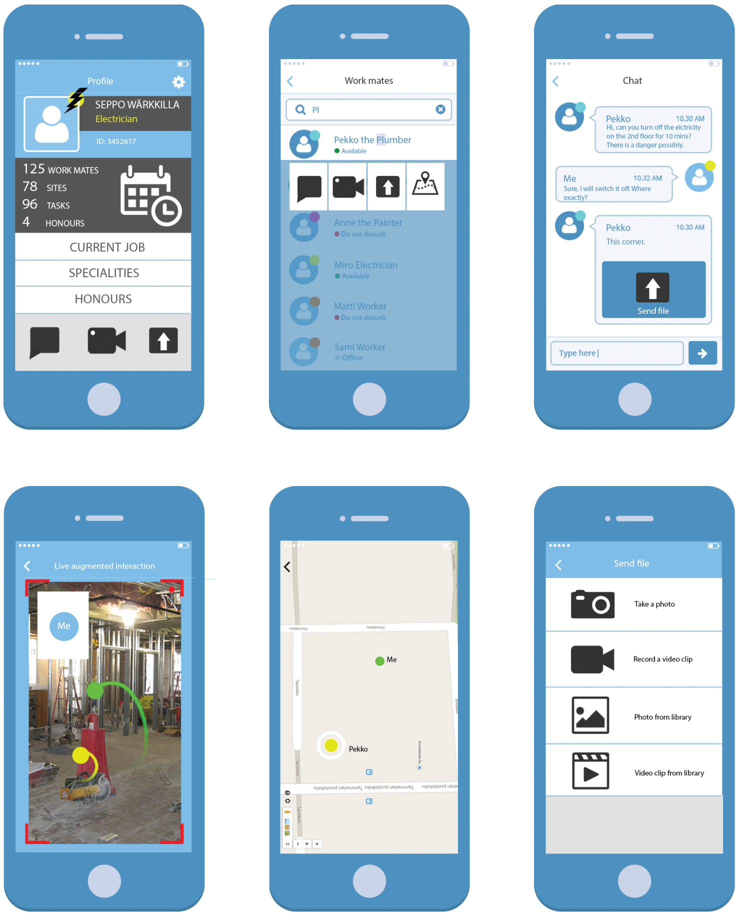



**Figure 3:** Mock-up views of the early SoAR concept by Jana Pejoska.

The proof-of-concept prototype of the vision-sharing feature was made in February 2015. The prototype was functional in a web browser over WebRTC protocol and it demonstrated augmented video conversations where the users could draw on the video stream and see each other's drawings. (Pejoska et al. 2016, pp. 479-480.)

## 3.3 Developing the Second Prototype

### *3.3.1 Product Design Phase: Technology choices*

In October 2015 LeGroup and I decided not to prototype the social networking features of the SoAR concept before the vision-sharing feature had been properly studied: vision-sharing was considered the priority to which the limited coding and design resources were assigned to. I started working full-time in the project in January 2015.

In Nov – Dec 2015 the second SoAR prototype was built with Apache Cordova in order to make effective use of Android native components (Figure 4). Apache Cordova is an open-source mobile development framework, which allows the use of standard web technologies (HTML5, CSS3, JavaScript) for cross-platform development (Apache Software Foundation 2016).

Android was chosen as the first mobile platform due to better WebRTC support, wider user base and its open-source basis. Web Real-Time Communication (WebRTC) handles the data streaming for SoAR. It is an open framework for browsers and web applications and it includes building blocks for web communication in the form of voice and video chat or peer-to-peer file sharing. Server technologies are required for managing the user registration and logging in, as well as initializing calls between users, but WebRTC data transmission takes place peer-to-peer, between browsers. WebRTC streaming and server connectivity demand a constant Wi-Fi or mobile network access. (WebRTC Initiative 2016.) The users need a phone number to sign up to SoAR, and they receive a confirmation code to this number via SMS.

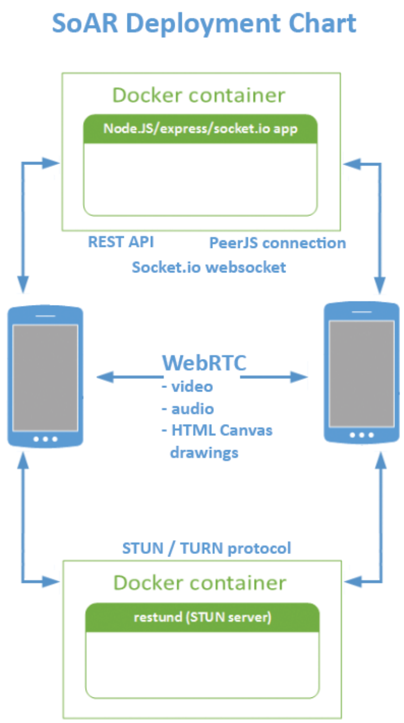



**Figure 4:** SoAR deployment chart by Matti Jokitulppo

Due to the open-source license, the source code of SoAR is available in the code repository GitHub and the application can be installed from the Android application store Google Play.[1] The technologies implemented in SoAR are described in more detail in Appendix 1.

### *3.3.2 SoAR and Depth-Sensing Technology*

As mobile devices are developing further and are even challenged by the development of head-mounted AR devices and other wearables (Carmigniani et al. 2010, Rolston 2013), it is important to assess whether designing augmented applications only for touch screen mobile phones and tablets is a suitable long-term solution. As the aim of the Learning Layers project was to develop digital tools that can immediately and affordably be taken to use by more than 1000 professional users (Learning Layers 2016), it was justified to concentrate mainly on currently common mobile platforms.

Regardless of the physical form of future communication devices, many of them will include depth-sensing technology for effective mapping, scanning and visualizing of the surroundings, as well as for motion and shape detection. Since the most commercially successful sensor Microsoft Kinect, released for console gaming in 2011, various companies have announced similar technologies that can also be used for mobile communication (Table 1).

---

[1] Github: https://github.com/learning-layers/sardroid
Google Play: https://goo.gl/Rvptn2

**Table 1:** Recent mobile depth sensing technologies

| Product | Company | Description | First released |
|---|---|---|---|
| Kinect | Microsoft | 3D sensor /console gaming | 2011 |
| Leap Motion | Leap Motion | Gesture detection sensor | 2012 |
| DUO | Code Laboratories | 3D sensor | 2013 |
| Structure Sensor | Occipital | 3D sensor | 2014 |
| Project Tango | Google | Computer vision platform | 2014 |
| RealSense | Intel | 3D sensor | 2015 |
| ZED | Stereolabs | 3D sensor | 2015 |
| Astra | Orbbec | 3D sensor | 2015 |
| Hololens | Microsoft | Head-mounted AR display | 2016 |

Because depth-sensing is on the verge of becoming a mainstream technology in mobile devices, the Intel RealSense 3D sensor was looked into as a possible prototyping device. The aim was to explore new kinds of augmentations to complement SoAR. Depth-sensing would make it possible to measure real-world distances between spots on a mobile screen, world-stabilizing the drawings of the users in the 3D space, focusing accurately on different depths in the view, 3D scanning of objects and testing on shape and motion recognition. According to Intel, the software development kit (SDK) for Android devices was to be released by the end of 2015. However, the SDK has been delayed and Intel appears to have prioritized moving ahead with Google Tango Project, the aim of which is to release a mobile device that combines the computer vision platform developed by Google with Intel's RealSense technology.

I wanted to ensure we had done everything in our power to get to utilize the RealSense developer assets and organized a meeting with Intel Research and Development team in Tampere, Feb 2016. However, Intel did not agree to allow the use of their unreleased products for research purposes. The setback led us to concentrate fully on testing and improving the drawing and vision-sharing features we could develop with the currently available mobile camera setup. Even if it would have been technically possible to make a depth-sensing SoAR prototype with bigger and less mobile 3D sensors and laptop computers, this would have diverted our efforts from the mobile development and we would have likely ended up with a lower quality application prototype while the coder resources would have had to be distributed to two different development platforms.

## 3.4 Usability Research

Before taking SoAR to actual work context it was imperative to conduct usability research in lab conditions to make sure the prototype would work as intended and actually advance collaboration in the workplace instead of being a disturbance. Apart from design needs, usability research also resulted in findings on the drawing feature and the nature of collaboration over video. These findings are compared to earlier research on similar applications in Chapter 4.

All usability research activities took place in the Media Lab of Aalto University, School of Arts, Design and Architecture. The users were Media lab students and research staff (Table 2).

**Table 2:** Usability research activities

| Activity | Users | Age, gender | Length | Time | Data format |
|---|---|---|---|---|---|
| Paper prototyping | 4 | 22-45, 1 female, 3 male | 10-18 min | 25th-26th Feb 2016 | Video, photos |
| Usability tests | 5 x 2 | 26-37, 6 female, 4 male | 11-25 min | 4th-10th Mar 2016 | Video |
| Interviews* | 4 + 5 x 2 = 14 | 22-45, 7 female, 7 male | 15-55 min | 25th Feb–10th Mar 2016 | Video |

* All users of the paper prototyping and usability tests were interviewed after the test.

When testing with users, I had to choose whether to emphasize scientific accuracy or iterative design, because the needs of the two are partially contradictory. In order to create a scientifically solid test setup, the number of users should have ideally been higher and the prototype app in the usability tests should have remained the same in each test situation. This approach yields a larger amount of quantitative data and results that are easier to compare than the results acquired with a small sample.

However, in their work on Comparative User Evaluation (CUE), Molich and Dumas (2008, p. 280) encourage developer teams to test with a small amount of users and iterate: "The motto should not be 'Five test participants are enough to find 70% of the problems', but 'Five test participants are enough to drive a useful, iterative process'". An excessive amount of test results may lead into severe productivity issues: even if the results were useful as such, the amount alone may not in every case justify the resources required for analysis.

I prioritized the design process and chose to use a relatively low number of users – four in paper prototyping and five pairs testing with the actual SoAR prototype on mobile devices. We also made changes to the application between iterations as soon as we found bugs or realized a feature was not working as well as it should.
It is clear from this point of view that our test setup was not as controlled as in earlier studies on video-mediated collaboration (see Chapter 2.2) but the trade-off is justified since our team was small, consisting only of me as the designer and Jokitulppo as the developer. Working as effectively as possible was important in order to move on to the most essential phase of the study, testing in the workplace.

To maximize the product improvement achieved from user tests, it is recommendable to involve developers in the whole testing process, from planning to analysis (Molich, Dumas 2008). The developer was present in the SoAR usability test occasions. This proved to be beneficial, and not only from the productivity perspective. In the course of the five user tests run at Media Lab, the developer started to take an active role: he would make remarks on the test situation during the interview, ask questions and sit closer to the users towards the end of the tests. When it came to teamwork, it was easier to justify the changes required in the design when both members of the team had testified the problems the users had faced. Holding on to the user-centric point of view and keeping common ground in general was easier when it was possible to refer to the common test experiences when developing SoAR further, instead of referring to spreadsheets or written documents only.

*3.4.1 Paper Prototyping*

Paper prototyping is a method where early product features are implemented in a low-fidelity prototype made of paper, which is then tested with real users. Paper prototyping allows cost-effective and quick testing of new features and their design implications, as new iterations do not require software development efforts. (Rettig 1994.)

SoAR was chosen as a group project on the MA level Interface Prototyping course I took in February 2016. I collaborated with four Media Lab students on a paper prototype of SoAR. We designed a tablet-sized paper replica of all the features that existed in the mobile prototype of SoAR to learn whether there were severe issues with the user interface (Figure 5). We also wanted to learn how adding user profiles and storing snapshots from the video conversations would affect the user experience and design – these were features that had been considered but have not yet been implemented in the digital SoAR prototype.

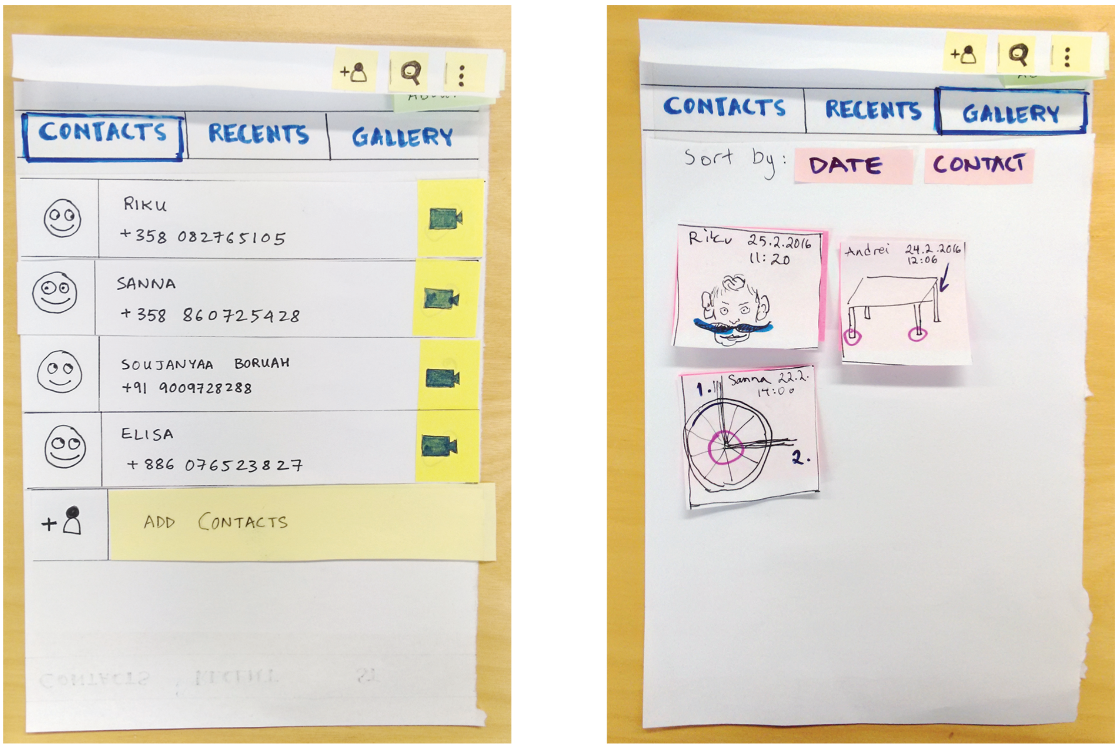



**Figure 5:** Some of the SoAR paper prototype UI views

The users completed a set of tasks where they used the paper prototype by clicking on the drawn buttons as if they were real ones on a mobile application. One of the team members rearranged the paper interface according to the user actions. The user tests were recorded on video.

We learned to conduct the tests more fluently along iterations and, following the paper prototyping routine, we also made quick changes and fixes in the design between iterations. This resulted in each user being able to finish the tasks faster than the previous one. Paper prototyping is thus not good for the most clinical, quantitative usability test approach because changes in the design undermine the comparability of much of the quantitative data.

Along with the achieved results, the paper prototyping experience was useful as a rehearsal before conducting the tests with the digital prototype. However, it was impossible to experience video conversations through a paper prototype: sitting in the same space with the conversation partner and drawing on a transparent in a paper prototype did not resemble actual remote collaboration.

### *3.4.2 Testing with the Digital Prototype*

The usability tests conducted with the digital prototype in March 2016 consisted of an introduction phase and a test task phase. The test devices were a Nexus 7 tablet and a Samsung S5 mobile phone and SoAR was used over the open Wi-Fi provided by Aalto University.

In the introduction phase, the user pairs were located in the same room with the developer and me. After a brief introduction to the project and the prototype, the users were instructed to register to SoAR, to add a contact and to make a test call with each other. During the test call, the users had to try the drawing feature and switching the views: initially the users saw their own camera view on top and the other person’s view below that. It was possible to make each of these views full-screen by touching the view (Figure 6).

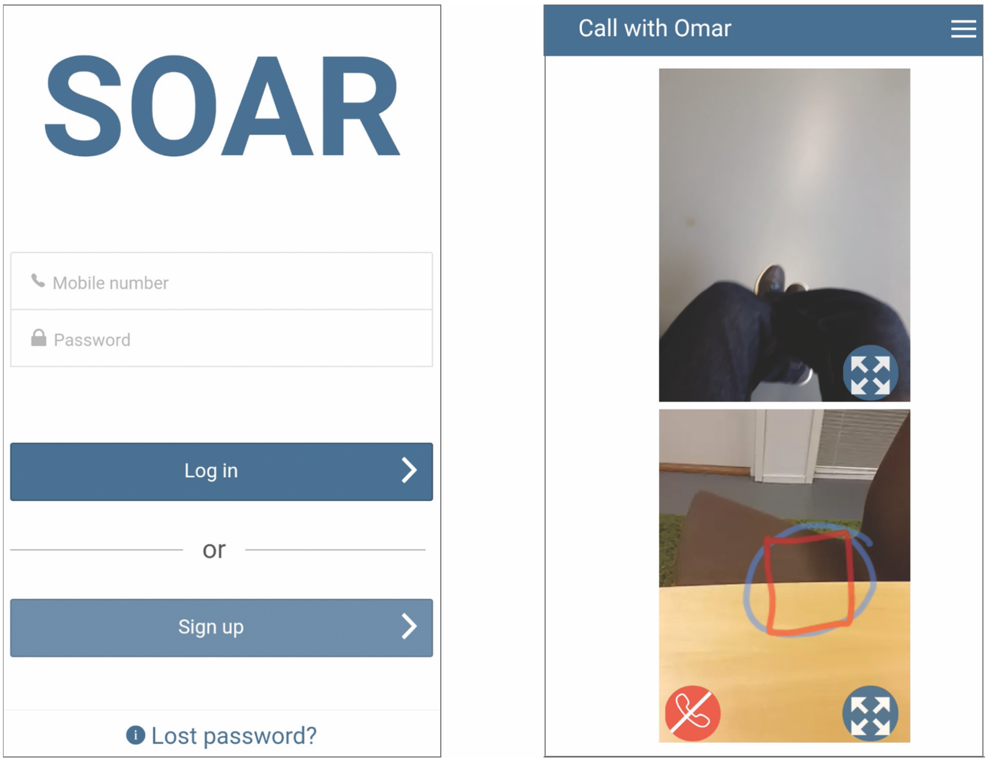



**Figure 6:** The login screen and the initial call view in usability test phase

For the task phase, the users determined which one of them was to be the *helper* and which one the *worker*. The helper stayed in the initial room as the worker moved to a classroom that had been prepared with the task objects. As the worker started a call with the helper, she was guided through a set of tasks by the helper who had the task instructions on a laptop.

The tasks included testing the drawing feature, finding a tray of Lego blocks hidden in the room, building a two-layer structure with the Lego bricks according to the helper's instructions, finding and arranging a set of characters from a large selection of similar items and drawing a picture on a whiteboard. The instruction was to solve the tasks with whatever means seemed the most convenient: the drawing feature was mentioned but not compulsory. Showing the helper's instructions directly from the screen over video was discouraged. The test tasks are included in Appendix 2.

The test sessions were recorded with two video recorders. One recorder was in the room where the helper worked and where the users also did the introductory part. I tested recording the screen events with a recording application installed on the mobile devices. This seemed to mess SoAR up, however. On Android devices only one application can currently use the microphone at any given time. Thus, as SoAR utilized the microphone, it was impossible to record sound. Also, there were more errors and connection issues when the screen recording software was on than when it was not utilized.

The usability issues were assessed based of severity to the user experience. The SoAR calls were transcribed and reviewed using conversation analysis methods: reviewing the discussions allowed me to understand how drawing on screen supplements spoken language. The factors I took into account were the length of the verbal instructions, the deictic reference types used both in conjunction and separately from the drawing gestures and references to directions, locations and task objects.

### *3.4.3 Interviews*

All 14 paper and digital prototype users were interviewed in semi-structured pair interviews after the test tasks. Interviews were necessary to ground the research and the design alike as the new knowledge in this thesis is manifested both in the digital prototype and the empirical data collected around it.

The data from the interviews was used for deepening the usability observations made from the video recordings of the tests. They were also necessary for revealing how the users experienced SoAR. According to Sampsa Hyysalo (2009, p. 52) technology is never a value in itself – a designer must do research to recognize the purposes the technology needs to serve. As products always have a personal dimension to them, it is important to take into account how the users feel about using the product. Feelings are essential to the adoption of new products.

It was also possible to tackle specific, technical and design issues in the prototype, as the usability test users were Media Lab students and staff who often had experience in designing digital products themselves. Because the test users did not represent the professional fields of construction, facility maintenance and quality management, their input on the usefulness of SoAR in the actual workplace gave guidance but could not be given such authority as the opinions collected from the field.

## 3.5 Testing in Work Context

Before initializing the field tests, there was a brief product design phase where the usability test findings were analyzed and bugs and design flaws were fixed. As a result, the in-call user interface went through a complete makeover (Figure 7). Because the decisions that led to the UI changes are based on and intertwined with the findings of the usability tests, the design changes are introduced in detail together with the analysis of the test outcomes in Chapter 4.7.


**Figure 7:** Call view in field test phase

I planned and conducted a set of research activities in work environments (Table 3), which I based on user-centered design principles described by Hyysalo (2009). As he writes, technologies are used in real situations and environments: using a technology consists of more than just performing the actions included in the mechanical use process. It is key to understand the ways the users are networked with each other, with other technologies and with the larger goals and tasks related to their work, along with how they feel about the technologies they use. It is considered a time and labor effective design practice to get acquainted with the line of work in question as the users can directly communicate their needs and expectations.

**Table 3:** Research activities conducted in work context

| Activity | Users | Age, gender | Time | Data format |
|---|---|---|---|---|
| Aalto – ISS field test | 2 | 35-45, 1 female, 1 male | 8th Feb – 9th Jun 2016 | Audio recordings, captured calls, photos |
| Skanska field test | 2 | 25-40, 2 male | 1st Mar - 3rd May 2016 | Audio recordings, captured calls, photos |
| BauABC field test | 17 | 16-40, 17 male | 1st Mar – 3rd May 2016 | Audio and video recordings, photos |
| Nokia Networks field test | 2 | 35-50, 2 male | 10th – 29th Jun 2016 | Audio recordings, captured calls |

I conducted semi-structured, themed interviews to map the types of work contexts where SoAR was introduced and the communication needs related to them. Final interviews were conducted after the test period to collect feedback of the usability and usefulness of the application in the line of work. Interviews related to the test participant's work have the risk of getting biased, as it can be difficult to be critical about one's own work-related practices in front of a researcher (Hyysalo 2009). Sometimes users rather describe how the work would ideally be, instead of how it is in reality.

To bridge the possible biases in the interview material and to get practical information about the day-to-day communication practices that are difficult to describe in an interview, I conducted job shadowing sessions. Shadowing is a form of field observation where the researcher follows an informant to conceive her actions as a whole. It is particularly useful when observing mobile work and mapping how workers interact and communicate (Hyysalo 2009). Apart from monitoring the course of the workday, I asked questions to get the participants explain to me what they were doing.

The material from the interviews and job shadowing sessions was analyzed thematically.

I aimed at following the same process with each field test case. Background interviews and job shadowing sessions were to be the beginning of each test period. The 5-day test period, during which the users were supposed to use SoAR independently, was to be followed by a final interview, where user experiences and further use ideas were to be collected (Figure 8).

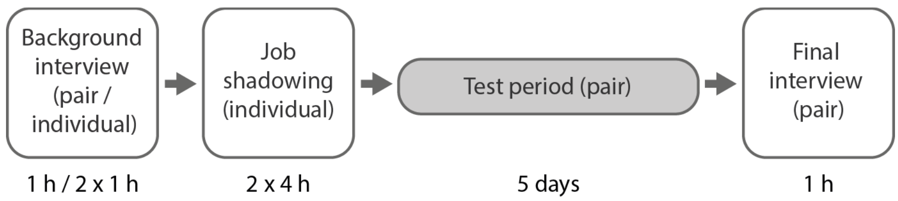



**Figure 8:** The ideal field test process

However, the work schedules of the test users and the Learning Layers project as well as the specifics of each test case forced me to make adjustments to the process.

Testing SoAR in actual work conditions proved to be challenging. As SoAR was meant to be a tool for help seeking and solving acute situations, it was impossible to predict all the circumstances where SoAR would be used. I could not be present at the workplaces constantly to observe the use. We implemented the possibility to record the calls in the app in order to store video and audio on the mobile devices for later reviewing. This feature was meant only for research use, as recording the calls of workers would likely result in suspicion and reluctance to use the app – if the calls could be recorded, it is also theoretically possible to use the recordings to evaluate the performance of the workers in acute, sometimes critical situations. When the app was recording SoAR calls, the frame rate of the video stream was lower than usual, which reduced the usability of the app. However, I considered this a bearable harm – the alternative documentation method would have been asking the participants to document the call situations by writing notes and taking photos. Writing is not as exact a way to document as recording, and this would have required extra work from the users.

As real-life work situations are often unpredictable, I quickly realized the 5-day test schedule was too short. Each of the field test periods conducted in actual workplaces was extended by at least five workdays. There were several reasons for the extended test periods, e.g. a participant got ill, the pair testing together noticed their work shifts allowed them only a limited, common work time to make calls, a participant forgot to charge the battery of the device or he forgot the device at home.

In general, as I was not present at the workplace when most of the field tests took place, I noticed it was necessary to check on the participants from time to time to encourage them to use the application. I also made SoAR calls with them during the test phase to make sure the app worked and they had a hang of using it. I could follow the call activity through the event analytics of the server and contact the test users if they appeared not to use SoAR. The content of the calls was reviewed in conjunction to related topics from the interviews

### *3.5.1 Skanska*

Skanska is an international construction and project development group. It is one of the largest companies operating in the Finnish construction sector. I conducted the field test with a construction surveyor and a foreman at Härmälänranta, Tampere. Härmälänranta is a new residential area, where Skanska's operations consisted of four different construction sites that were located side by side. The test users were recruited together with the shop steward and the development manager of Skanska Pirkanmaa, who considered SoAR a useful communication tool for foremen and surveyors.

A construction surveyor ensures that all structures at a construction site are built on the exact locations specified in the construction plan. Surveyor's tasks include staking out reference points at the site and taking correction measurements using a total station in conjunction with modeling software. The surveyor who tested SoAR was working at a site that was in the foundation phase, and his tasks included measuring and staking out the locations of the footing of the building. He worked in close quarters with a team of carpenters and ironworkers who built the footing molds.

Construction foremen are responsible for the division of labor and the overall progress of the construction work. They supervise subcontractors and occupational safety, and communicate with construction planners and the superiors in the main contractor company. The sites at Härmälänranta, Tampere had 2-3 foremen, depending on the working phase. The foreman who tested SoAR was supervising the interior phase of one of the apartment blocks during the test period.

The field test activities took place between 1st March and 3rd May 2016 (Figure 9). I recorded the background interview and took notes and photos during job shadowing. The final interview was recorded in writing.

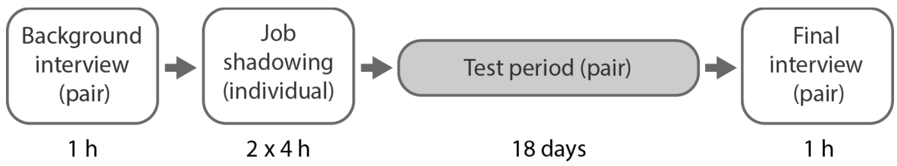



**Figure 9:** Skanska field test activities

The background interview was conducted as a pair interview and even if this is generally a good way to learn about a new line of work, the discussion may have been biased at times, as the foreman was a superior of the surveyor. Skanska's shop steward was also present and took part in the discussion. It is impossible to tell afterwards how the data would have been different had the participants been interviewed separately. On the other hand, the workers had a different point of view to communication at a construction site, which added to the conversation.

Eventually, the biggest challenge of this test case was that the foreman and surveyor ended up working on different sites of Härmälänranta and had no real collaboration situations. The conditions of using SoAR were real as the test calls were made at the site, but the work situations related to the calls were simulated. The test period was extended to 18 working days partly due to the different work locations and scheduling difficulties that resulted from this. There were also complications with recording the SoAR calls: all of the calls could not be retrieved afterwards. The bug was fixed before the next field tests.

The test devices were two Samsung S5 phones: one of them belonged to LeGroup and the other one to the surveyor who preferred to use his own device for testing. The calls were made over mobile network as there was no Wi-Fi coverage over all the Härmälänranta sites.

### *3.5.2 Aalto Campus Services and ISS Facility Maintenance*

Aalto Campus Services is responsible for the services related to the infrastructure and facilities of Aalto University. ISS Palvelut is the Finnish branch of the international ISS Group, which offers a range of facility services. The field test was conducted with an Aalto Campus Services lobby attendant located at Miestentie 3, Espoo, and an ISS maintenance person who was responsible for the Miestentie 3 facility.

Lobby attendants at Miestentie 3 serve primarily the users of the building, the staff and the students of the Departments of Architecture and Media, by monitoring accessibility, safety and the condition of the facility and its movables. They provide visitor guidance and assistance with presentation and printing technology and space reservations, as well as promote energy efficient use of the facility.

Facility maintenance personnel at ISS are typically responsible for taking care of the service calls and routine tasks of a selection of facilities located close to each other. The routine tasks include e.g. reparations, taking care of the outdoor areas, taking monthly electricity and water readings and doing rounds in the facilities. Maintenance personnel get service calls through several facility management systems and they also operate on the basis of the information received from different building automation systems. Often maintenance persons also work on call for additional service areas.

I interviewed the lobby attendant who agreed to test SoAR and chose the facility maintenance person as the second test user on the basis of the background interview.

It appeared the maintenance person was an optimal test partner as there already was an established communication pattern for resolving the issues at the Miestentie 3 facility. The test activities took place between 8th February and 9th June 2016 (Figure 10). The interviews were mostly recorded, apart from some hand-written notes. The job shadowing sessions were documented with photos and notes.

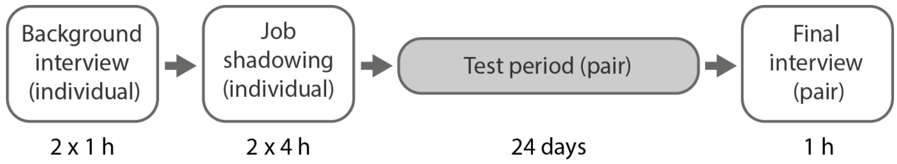



**Figure 10:** Aalto Campus Services and ISS Facility Maintenance field test activities

There were some technical and practical complications during the test period. Earlier tests at Skanska and Media Lab had not uncovered an issue with logging in to the server that almost disrupted the tests. SoAR would also at times show a user as "offline" and thus not available for calls if the application was not actively used for a while, which was likely due to the native functionality of the test phones. The malfunctions were demotivating, as the participants were looking forward to solving real faults in the facility with SoAR. On the other hand, such a long-term use period was the best way to uncover such bugs: debugging and development work was more fast-paced and involved lots of periodic, short calls that were not enough to reveal them. I tried to help the situation by extending the test period and providing instructions on how to recover from malfunctions.

The test devices were two Samsung S5 phones from LeGroup. The calls were made over the Aalto University open Wi-Fi and mobile network.

### *3.5.3 BauABC Rostrup*

BauABC Rostrup is one of the largest construction training centers in Germany. Apprentices of various building occupations take part in initial training before moving on to learning on the job in construction companies, and return for further training periods before graduating to the profession. BauABC Rostrup was one of the industry partners in the Learning Layers project, and the SoAR field test was a part of the piloting program of all the digital tools produced in the project.

The operator apprentices at BauABC were a group of 17, with different backgrounds. They work in companies mostly in the area of Lower Saxony, Germany and return to the BauABC training facility regularly to complete sets of exercises with different construction machines to ensure they meet the level of expertise required on the job.

The field test process was considerably different from the other test cases, as I only had the chance to test SoAR at BauABC on one day, 25th May 2016 (Figure 11). This case took place in a practical education setting instead of actual work context, and it was fully facilitated whereas the other users used the application independently.

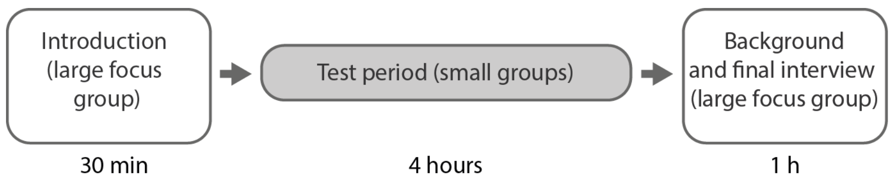



**Figure 11:** BauABC field test activities

The test session consisted of a short introduction of how SoAR works, the actual field test in small groups and a focus group interview about the communication practices at work and the user experience of SoAR. The interview was recorded and the field test was documented with video recordings, written notes and photos. The calls were not recorded on the devices because I was able to be present when the calls were made.

The apprentices were working in small groups spread apart from each other in the facility grounds, and their tasks included for example using excavators and installing sheet piling into the ground. The instructor of the group had one of the test devices while I went from one small group to another and encouraged the apprentices to make SoAR calls with the instructor to ask for advice if needed or just to report how they were doing with the tasks. The apprentices were positive towards testing SoAR and the setting was useful in general: the instructor had to move a lot around the facility and the apprentices had thus an actual need to collaborate remotely when he was not in the vicinity.

BauABC case was the only field test where the developer was present. In BauABC, the emphasis was on using the application, whereas the other field tests took place over several weeks and involved more interviews and free-form discussions than working with the actual prototype – these field tests would not have been effective use of the developer's time. The time available for testing in BauABC was limited and the effort to organize the session in Germany was considerable. In order to maximize the chance to succeed it was crucial to have the developer on site to respond to acute technical issues and to gather mutual experiences on the usefulness of the app in actual work context within a practical time frame.

The test devices were two Samsung S5 phones from LeGroup, and the calls were made over the mobile network.

### *3.5.4 Nokia*

Nokia is among the world-leading companies in network infrastructure technologies. I conducted the field test with two quality managers who agreed to simulate how SoAR could be used to support remote workshops and internal audits.

Quality managers are responsible for maintaining and developing processes and the quality management system of the company, in compliance to standards. This requires audits and inspections, making and ordering reports, organizing training sessions and making for example process descriptions and manuals. Quality managers look into the whole supply chain of products and services, for Nokia corporation and its external suppliers alike. They also lead development projects.

The test activities took place between 10th and 29th June 2016 (Figure 12). The interviews and SoAR calls were recorded.


**Figure 12:** Nokia field test activities

The possibility to test with Nokia came up late in the project, so the field test was slightly narrowed down compared to the Skanska and Aalto–ISS cases. Job shadowing was left out because it was not possible to follow a real auditing case in an offshore location. The test activities took place in Nokia headquarters in Espoo apart from one test call, which was made from an internal workshop in Germany to the local colleagues in Finland.

The users suffered from connection issues during the test phase – they reported several problems with disconnected calls and lack of audio or video during the calls. The users were keen on introducing SoAR to collect ideas and opinions from other workers at Nokia as well, but the malfunctions hampered with some of these opportunities.

The test devices were two Samsung S5 phones from LeGroup, and the calls were made over mobile network.

# 4. REMOTE COLLABORATION IN A CONTROLLED ENVIRONMENT

The key results of this chapter are based on the usability tests conducted with the digital prototype of SoAR. The collaborating partners in the usability tests are called *helper* and *worker.* The worker must accomplish tasks that the helper describes remotely over a SoAR call. These terms are chosen only for the sake of clarity only; in real life such clear-cut roles obviously do not apply since help seeking, giving advice, and task-oriented performance blend in most occupations. The test cases are numbered from 1 to 5, e.g. Helper 2 (short H2) refers to the helper from test case 2. The test setup is described in more detail in Chapter 3.4.2 and the actual tasks are documented in Appendix 2.

The usability test tasks were designed to resemble physical work in a sense that they required mobility and different types of actions: searching and moving objects in a room, identifying and selecting objects from a group of similar ones, assembling a structure and drawing. Due to the variety of tasks there were many approaches available for the users and they were instructed to use whatever methods they wanted to solve the tasks. Unlike in the earlier studies described in Chapter 2.3.2, drawing on the screen was not mandatory. I wanted to monitor what the users chose to do instead of only mapping their preferences in interviews: expressed preference does not necessarily equal actual adoption and use of the feature.

Task performance times and error rates have been analyzed in many studies on shared view and augmented video collaboration, typically in comparison to non-augmented or audio-only technologies. E.g. Fussell et al. (2004, p. 296) found out that using an augmented pointing system together with video connection allows helpers to guide workers through a test task using fewer words. The results of the usability tests of SoAR were not consistent in this sense due to design changes in the prototype, the small amount of users and changes in the tasks and task order (Table 4). The users were not encouraged to perform as fast as possible and small errors in task performance were not taken into account: more emphasis was placed on perceived advantages or difficulties in using SoAR. Along with observations on user behavior, I also looked into communication patterns that were discernible in the transcribed conversations of the users.

**Table 4:** Helper word count for the most verbally demanding tasks (Red = drawing on screen was not utilized)

| | Helper 1 | Helper 2 | Helper 3 | Helper 4 | Helper 5 |
|---|---|---|---|---|---|
| Pick up the right Lego bricks | 200 | 77 | 128 | 123 | 138 |
| Build Lego layer 1 | 220 | 146 | 154 | 183* | 302 |
| Build Lego layer 2 | 138 | 108 | 103 | 113* | 136 |
| Choose the right four characters | 263 | 127 | 192 | 199** | 132 |
| Draw a face on a whiteboard | 404 | 132 | 165 | 60 | 111 |
| total | 1225 | 590 | 742 | 678 | 819 |

* No audio from the worker to helper due to a streaming bug

** Decided to instruct mainly by drawing, not referring to colors and shapes.

Helper word count was not necessarily a suitable measure of success in such a flexible task setting. Regardless of how much the pairs utilized the drawing feature, helpers and workers understood a task differently at times. Recovering from such misunderstandings required the helper to explain what went wrong and how to track back, and this in turn added to the word count. Misunderstandings cannot be fully avoided in collaboration, even if a larger user sample might shed light on whether drawing on screen has the capability to reduce them. Also, as there was no time pressure, personal speaking habits of the helpers seemed to affect the length of the utterances. Some were for example polite: "There… that 7. Could you pick it up?" (helper 5) whereas others used more compact language in the same task: " Ok, so… this one… err, this… ok, this one." (helper 4.)

There were also some bugs in the prototype, which changed user behavior. Several test pairs suffered from disconnected calls. A recurring issue with WebRTC was that one or more of the streamed media elements would not come through. During the usability tests, drawings got briefly disabled during test case 5. In case 4, audio from the worker did not come through, but the pair was able to finish all the tasks despite the bug.

## 4.1 Screen Confusions

The SoAR version used in the tests opened the camera views of both users in equal size when a new call was initiated. The users could choose whether they would work with both views, their own view or the remote view.

Despite large “fullscreen” icons, users often stayed in the default two-view formation, sacrificing thus a portion of the task space view (Figure 13). There was no particular benefit in working with the helper view: the helper’s camera mostly pointed to a table or a wall in the room she was in. Helpers typically switched to fullscreen when they noticed it was difficult to draw on a small screen. All in all, there were no clear patterns in which view mode the users worked. When drawing, the users utilized mostly the worker’s camera view.

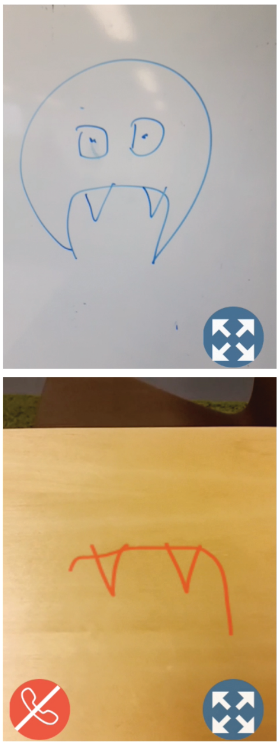


**Figure 13:** Helper 4 drawing on his own screen, worker 4 on a whiteboard

It was difficult to determine which screen the conversation partner was working on with this view arrangement. Possible solutions envisioned by the users were controlling remotely which view the partner sees (helper 3), or highlighting which screen is active on the partner's device (helper 4).

Interestingly, the users reported initial difficulties with recognizing which view was their own. Rear camera view is probably not as intuitively recognized as the front camera, which allows facial view. The users often started the calls by confirming the call status with questions and statements like "are you there" and "I can hear you", even if they could see that the video connection was established. There were no facial expressions available for recognition and no conventions to resort to – the customary ways to initiate a regular audio call apparently did not apply well for SoAR calls.

Workers 1 and 4 suggested the front camera should be default when a SoAR call is initiated, after which the camera could be flipped manually. I decided to collect feedback from industry users before making any decisions on the matter: I found it necessary to assess whether it was more important to enable concentrating on the work environment as efficiently as possible through the rear camera, or whether the screen confusions and discomfort were so severe that starting a call with the front view was necessary. The possibility to flip the view later might not be evident to the users, either, as most video conversations are conducted in the “talking heads” formation only. Likely these issues will turn obsolete if 360° cameras for mobile devices become more commonplace in the near future.

## 4.2 The Choice to Draw or Not

All users expressed in themed interviews that they were positive about the usefulness SoAR: having a shared view and being able to draw on it were considered beneficial for remote collaboration. All test pairs made a rehearsal call with each other before the actual test tasks to ensure they were familiar with the drawing feature.

However, even if the feature was simple to use – it only required touching the screen during the call - and the users had just tried it out, the tasks apparently demanded such an amount of concentration that the drawing feature was easy to forget. Test pairs 1 and 3 completely forgot about the drawing feature as soon as the task phase started. In my opinion this was only because drawing on screen was new to the users in this context, not because they would have objected the feature or felt that other methods to solve the tasks were more efficient. Helper 3 expressed surprise when she realized she had not thought of drawing:

> H3: Yeah. And now we have a, umm, black [character] five on a white background. I can see it.
> W3: Yeah?
> H3: It is, aa, on the left, um, south of where you picked up the [character] P.
> W3: Just, just draw on the screen if you can see it.
> H3: yeah, it's... ah ok! Yeah, yeah, I didn't think of that (laughter).

Helper 3 gestured a lot at the mobile screen while instructing the worker verbally: she was even pointing at the right characters with a finger which I interpret as frustration of not having – or remembering – a means to point out the exact task objects to the worker. Also helper 1 explained in the interview that drawing did not even come to her mind until she simply could not do without it when she had to instruct the worker how to draw a face on a whiteboard.

Adopting new problem-solving methods is difficult under high cognitive load. If there is not enough time to prepare and adjust to using a new tactic, it is easy to resort to familiar methods that have been found useful in the past. Apparently the load from the test situation was enough for some helpers to resort to describing features of the task objects instead of drawing, which I could not properly anticipate.

In response, I modified the task order and added drawing rehearsals to the beginning after test case 1. Additional practice did not make a difference for Helper 3, who still forgot about drawing. For test cases 4 and 5 I also moved the whiteboard task to the beginning because helpers 1-3 had proven this task was intuitive to accomplish utilizing drawing on screen.

Only helpers made use of the drawing feature during the task phase, save for worker 4 who used drawings to communicate he could still hear the helper even if his own voice did not go through. Since the workers were concentrating on physical tasks, they appeared to have a strong preference to work with the task objects in the real world, not mediated through the screen. Pointing at the actual objects instead of pointing by drawing seems to have been more intuitive for all workers. Worker 4 explained the perspective on the screen was so different from the live view it took him time to adjust to looking back at the screen again.

The workers were responsible for keeping the helpers aware of the physical surroundings of the tasks, and they came up with several ways to highlight the details and allow helpers to judge whether the tasks were conducted correctly. All workers brought the camera closer to the task objects and lifted items closer to the camera. They also needed to do stabilizing maneuvers with the mobile device to keep the view still, because the drawings of the helpers did not stick to details in this prototype version. Workers 1, 3, 4 and 5 pointed with a finger at task objects to assist with choosing the right ones, and worker 3 also arranged similar objects side by side for comparison. Worker 4 spread Lego bricks apart so the helper could easily point at them by drawing.

## 4.3 Overview of the Drawing Gestures

Almost all the drawing gestures done by the helpers were circles aimed at pointing out items and locations. This is in line with the findings of Fussell et al. (2004, p. 299) – according to their study on drawing in video-mediated remote collaboration, a clear majority (75 %) of drawings done over video view consisted of pointing at objects and locations. The other types recorded were drawings that pointed out directions and orientation (less than 20 %) and sketches. Premade, round pointers have been found to be less efficient for task performance (Kim et al. 2013, Fussell et al. 2004).

Helpers 2, 4 and 5 also drew directions and orientations during the Lego construction tasks. Helper 5 drew the outlines and orientations of each block, whereas helpers 2 and 4 quickly drafted the rough alignment of each brick. The latter two also started the task by outlining the final shape of the structure (Figure 14). Worker 2 found this disturbing, but worker 4 stated he benefitted from the outline drawing.

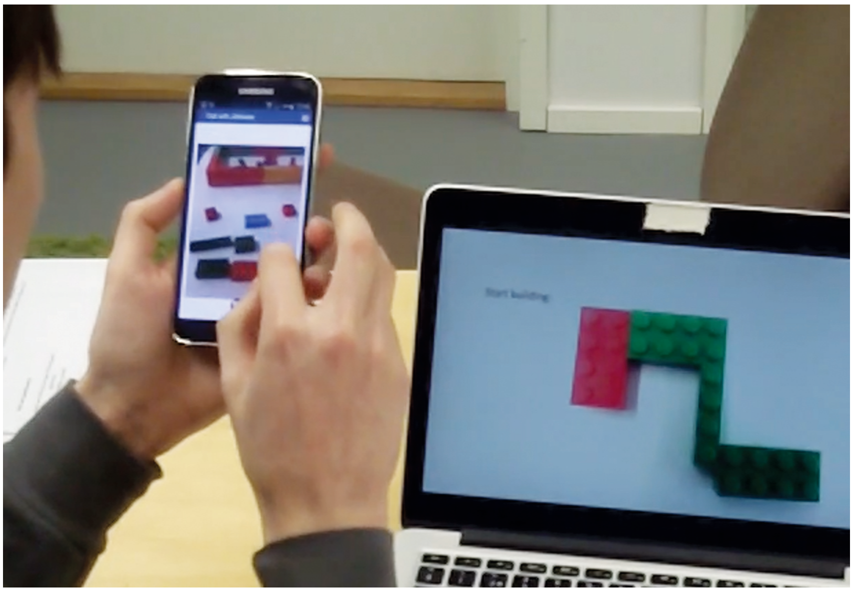


**Figure 14:** Helper 4 drawing outlines of the structure (video screenshot)

Automatically erased annotations have been found to speed up collaboration in some assembly tasks (Ou et al. 2003, Fussell et al. 2004). If the task procedure consists of a series of similar, stationary work phases, users likely learn to sync their actions, and auto-erase may speed up the work. Real-life work is seldom as predictable and mechanical, however, and giving the users control over erasing the annotations may lead into a smaller number of mistakes (Kim et al. 2013).

In SoAR usability tests, the drawings were erased automatically after 2 seconds, which the users found confusing. Due to negative feedback we increased the auto-erase timeout to 6 seconds before the test case 4, but there was no noticeable improvement in user experience. All users objected auto-erase in all tasks, but most prominently in the face-drawing task when they needed to make exact drawings. It seems clear that auto-erase is not a useful feature, should SoAR be used in any real work context. Manual erase function would likely have also improved the performance of those test pairs that utilized drawing by reducing the need to repeat instructions.

Drawing gestures did not seem to fully replace speech at any point. The importance of drawing gestures can only be evaluated in conjunction with speech: how the users support their message by drawing, what they say and what they can leave out due to drawings.

## 4.4 Conversational Characteristics

In order to better understand how common ground is built, Fussell et al. (2000) discussed the utterances in collaborative situations utilizing content categories. Each of the content types can appear as a question, answer or statement. Individual utterances from SoAR test cases seem to fit this classification (Table 5).

**Table 5:** Classification of utterances

| **Content category** | **Examples from SoAR tests** |
|---|---|
| *Acknowledgement:* confirms the message has been understood | yeah; mm; ok |
| *Internal state:* intentions, knowledge and emotions | I didn't think of doing that; this is quite tricky |
| *Task status:* state of the task and the task objects | I think we're done; we're missing one piece |
| *Procedural:* instructions furthering completion | you need to pick up the biggest red; can you show me the cupboard in the room? |
| *Referential:* advance identification of objects and locations | this black thing? ; it's in the center of your screen, it's yellow background with red character P |

However, when these categories are applied in actual conversations that resemble real life discussions, the amount of utterances makes analysis fairly complex (Table 6).

**Table 6:** Content categories in Lego picking task (underlined parts occur together with drawing gestures)

| | | **Content category** |
|---|---|---|
| H2: | Ok, so, now you pick, ah... | Procedural |
| | whoops, you're moving so much. | Internal state |
| | This block, the long green one. | Referential / procedural |
| W2: | (Picks a correct brick) | |
| H2: | Then... one of these, short green, | Referential / procedural |
| W2: | (Picks a correct brick) | |
| H2: | And, the other short green. | Referential / procedural |
| W2: | (Picks a correct brick) | |
| H2: | And... one, like this red | Referential / procedural |
| W2: | (Picks a wrong brick) | |
| H2: | No no, | Task status |
| | the other red, | Referential / procedural |
| | No no… | Task status |
| W2: | (Picks a correct brick) | |
| H2: | Yes ok, this one, | Referential / task status |
| | and the... two square reds. This and this. | Referential / procedural |
| | There's something wrong with the drawings because they don't disappear… | Internal state |

| | | |
|---|---|---|
| | … here. | Referential / procedural |
| W2: | (Picks correct bricks) | |
| H2: | And the blue one, the dark blue one, this one, | Referential / procedural |
| W2: | This one, and... | Referential / task status |
| H2: | Yes. | Task status |
| W2: | … what | Task status |
| H2: | And that's it. | Task status |

All in all, a communication pattern can still be discerned here: an initial procedural utterance leads into common ground, which is then challenged by a false choice by the worker. The status of the task is assessed and corrected. Finally, the users conclude whether the task has reached the final state. It is noteworthy that while shared view and pointing by drawing are available, procedural instructions and sometimes also statements about the task status melt into referential utterances: "This one" refers to a task object and simultaneously means it should be picked up. "Yes, ok, this one" consists of confirmation that the already chosen brick is the right one – reference underlines the choice. These melt categories also often occur together with drawing gestures, which could indicate how drawing makes communication more effective: it bridges two communication needs into one utterance. As drawing on screen clearly occurs together with referential utterances, I will give referencing a closer look in Chapter 4.5.

The workers communicated verbally much less than the helpers, as expected: the helpers were responsible of almost all procedural utterances, which were significantly longer than the other types. The workers' communication consisted of mainly of short acknowledgements (yeah, ok, mm, alright, cool) and short clarifying questions about the task status or reference (Like this? Do I attach? Is that right? This one?).

As in the example above, in all test cases there were times when the helper went on without needing a verbal acknowledgement from the worker. Seeing that the worker had performed the correct action was enough to maintain situational awareness - the same notion has been made by e.g. Gergle et al. (2004). However, the helpers typically kept making constant task status confirmations, like stating "yes" or "ok" after each correct action made by the worker.

The workers could also often anticipate what the helpers meant and acted accordingly before the instructions were finished:

> H1: And then the... next one will be the same, as...
> W1: (picks the right brick)
> H1: yes. And then there is one left, that is a blue one, it's with eight studs. Err, but, yes, in two...
> W1: (shows two bricks to the camera)

Common ground manifested also as effective adjustments of the shared view, done by workers. As helper 5 describes to worker 5: “It takes a lot of collaboration, like, you understood what I would need to see in order for me to show you what you needed to do. So, when you were naturally kind of zooming in without me saying anything, I liked that.” The notion also underlines that sense of connection plays a role in a collaborative task. Some pairs were getting along visibly well: they were relaxed and joking while others seemed more timid. Some helpers struggled to make workers comprehend what the tasks were all about. Even if the user sample is too small to assess the effect of the relationship chemistry of the test pairs, it seems safe to say the pairs who were getting along well likely benefitted of this.

## 4.5 Referencing

Earlier studies have shown that collaborating in shared view allows efficient referencing to objects and locations using deictic references (deixes): *this, these, here, that, those* and *there* (Bauer et al. 1999, Kraut et al. 2003, Fussell et al. 2004). Typically deictic referencing takes place simultaneously with gestures and supports them (Bauer et al. 1999). *This, these* and *here* are often used in co-located situations, whereas *that, those* and *there* suggest there is distance between collaborators. When pointing is available in remote collaboration, local deixes are used more often (Fussell et al. 2004).

The users had to refer to objects and locations most often in the Lego tasks and in the task in which they had to pick up characters (Figure 15). I coded the local and remote deixes in the users’ conversations during these tasks, as well as direct references to the features of the task objects (red, square, character P) and spatial and directional references (90 degrees, in line, next to). Both direct and spatial references can occur independently (a red brick, horizontal) or in relation to other objects or the task space (the bigger red brick, on the right side of).

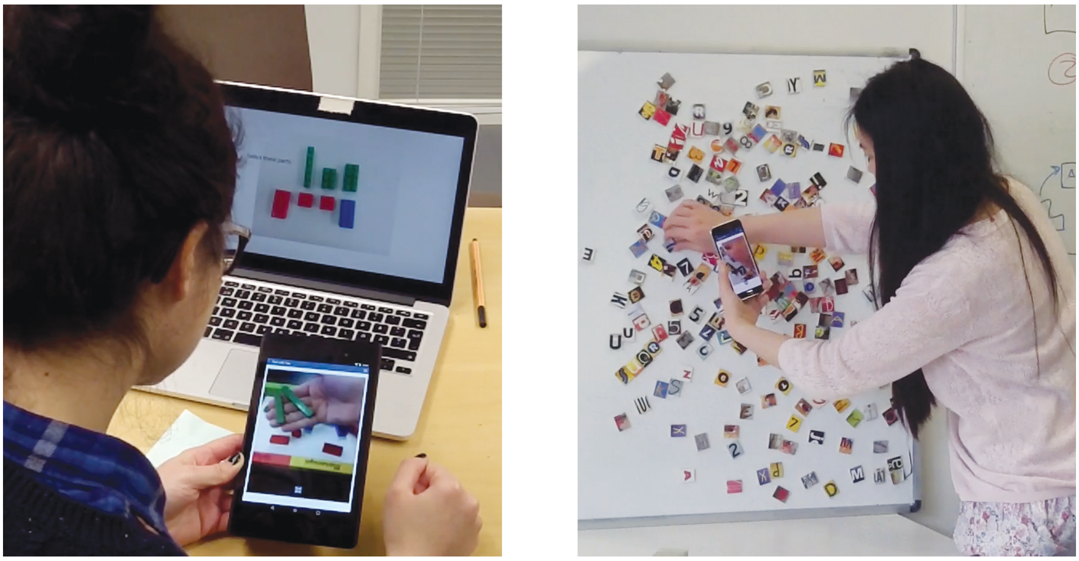


**Figure 15**: Picking up Legos (helper 3) and characters (worker 1)

When choosing the correct items from a group of similar objects, helpers who utilized drawing used significantly less direct feature reference than those who did not draw. Local deixes that occurred simultaneously with drawing gestures were common: “Pick this one”. While drawing, it was also typical to speak out the primary feature of the object to allow easy identification – name of the character or, with Legos, the color of the brick: “this 5”; “this p here”; “this red one”. The drawing helpers mentioned a second feature - the color of the letter or the shape of the Lego brick - only occasionally and always in connection with the primary feature. Spatial instructions were practically not needed at all: the location of the task object was only mentioned once.

The helpers who did not draw had to resort to lengthy descriptions of the features. They used mostly remote deixis “that”, but also occasionally local deixis “this”. Picking up the correct characters required using spatial instructions, whereas the amount of

Lego bricks from which to select was smaller and spatial references were hardly needed at all.

> H3: Now, there's a P letter. Umm, it is red and on a yellow background. I see it, it's somewhere in the middle.
>
> H1: --- a blue one, it's with eight studs. Err, but, yes, in two [columns]… err… the darker blue one.

All workers used predominantly local deixes. They apparently experienced no shift in attention between local and remote targets like the helpers did, as the latter seemed to consider the items and locations they were able to point at themselves as local and the rest as remote.

Without the drawings there appeared to be much more misunderstandings that had to be corrected with more detailed descriptions. Shared view was an advantage, however: workers were able to physically point at items and helpers could correct them quickly. According to the users it was fairly easy to collaborate without drawing as well, even if the utterances seem clumsy.

Interestingly, helper 4 attempted to instruct the worker without describing task objects at all, just by speaking in general terms and drawing on the screen. Even if this method was a result of a misunderstanding – helper 4 did not initially realize he was supposed to use the most convenient methods, not only drawing – it shed light on the importance of descriptive terms in collaboration. Helper 4 quickly grew frustrated in verbally complex tasks and eventually switched to describing task objects after several misunderstandings and mistakes:

> H4: The character or letters, yeah, this... I'll circle those that you shall then put… err... beneath the picture you just drew. Ok, so... <u>this one</u>... err, this... ok, <u>this one</u>...
> W4: Yeah?
> H4: And <u>this</u>.
> W4: Ok what is the other one?
> H4: Ok, <u>this one</u>.
> W4: And... err... this?

H4: No no no no errr... that.

W4: Number two?

H4: No no... umm... ok. Can you see the one, now? Now it's in the circle.

W4: This black thing?

H4: No no no. Umm. Oh the one just... I just use the feature of the app instead saying the character but... yeah.

(Underlined words occur simultaneously with drawing on screen.)

According to related research, referring to angles and directions verbally is more difficult than describing task objects (Fussell et al. 2000, Fussell et al. 2003, Gauglitz et al. 2014b). SoAR usability tests clearly supported the claim. Confusions appeared especially when helpers had to describe directions in 3D space, because terms like up, above and on top are ambiguous when working on a 2D screen: “On top of them. Like… on… top. Not above, but… you know, in 3D space, you put it on top [laughter]” (Helper 2, building with Lego bricks). Drawing on screen did not remove this problem, which seemed to appear arbitrarily, independent of the methods the pair was using to solve the tasks.

Describing a shape that deviated from previous ones was also challenging. Placing a Lego brick that was aligned differently from the others required several attempts from all user pairs (Figure 16).

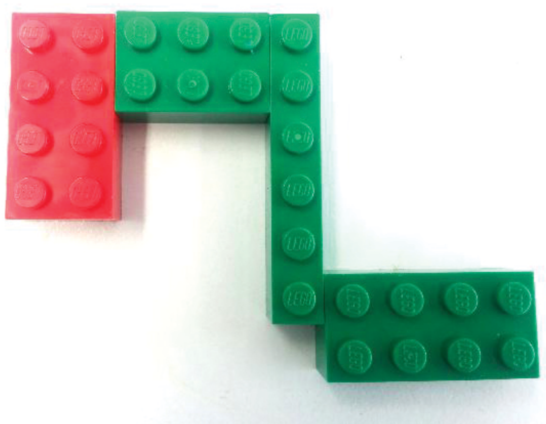


**Figure 16:** The difficult alignment of the last green block

Drawing the exact position of the different alignment on the screen was difficult with the SoAR prototype because of non-stabilized drawings. A common method was to first have the piece placed on the right side of the structure and guiding it little by little to the right spot by continuous instructions: “keep going, keep going” or “down, down”.

## 4.6 Discussion: Towards Testing in Actual Work Conditions

Bauer et al. (1999, p. 5) concluded in their study on collaboration over an augmented video-call system that pointing was the decisive way to instruct a remote worker, more important than speech alone or the combination of the two. Pointing significantly reduced the amount of verbal instructions and half of the helpers hardly used any spoken cues. The tasks in this study were stationary and rigid, however, and the users likely learned to utilize a communication pattern suitable for the specific tasks in question.

Even if the results from SoAR tests are not comprehensive, I claim that in real life pointing and speech are needed alike in remote collaboration: as soon as the worker needs to move or switch from one type of a task to another, the need for verbal communication builds up. However, the helpers who resorted to drawing, would primarily attempt at instructing by drawing and related deixes, and they mostly referred to the features of the task objects only after the pointing gesture. The word count of the helpers did not prove drawing more effective than video-only communication. Zooming in on the exact references that the helpers made shows, however, that pointing by drawing allows significantly more compact expressions than collaboration over video. The helpers used descriptive references along with deixes most of the time to maintain situational awareness.

All in all, the usability tests support earlier findings (Chapter 2.2.2) on the benefits of augmented video stream in collaboration: users experienced drawing on screen as preferable to utilizing only video and audio, and the instructions of the helpers were more compact and exact when drawing. I however find that static assembly tasks and numerical analyses of most of the earlier studies are not adequate for developing a tool meant for physical, mobile work. Small misunderstandings, unintentional movements or differences between the communication habits of the users have a substantial effect on numerical variables like the word count or completion time, when the tasks are versatile and require mobility. These effects are typical of real-life work and will not necessarily surface in a static task setting.

Drawing gestures were only used by the helpers. Even so, it is important that the drawing possibility remains available to both users. As noted earlier, *helper* and *worker* are not functional terms in actual work context, because both helping and seeking help are required in most lines of work, even if there may be established collaborative patterns between specific workers. Reaching situation awareness in any real collaborative situation calls for active participation and verbal assessment. A helper or an expert would not be just a repository of information like in the SoAR test setup. A worker or a help-seeker could not only wait for instructions - instead, she would actively communicate the situation to the helper and explain emerging details along the way.

It is clear that many of the issues the users had were due to flaws in the prototype design. Improvements in the prototype are likely to enable more effective discussion and reference to task items and locations, as well as remove some of the strain experienced by the users when solving tasks with SoAR.

## 4.7 Design Implications

The paper prototyping sessions provided with some food of thought when it comes to making new design choices, even if the findings did not prove quite as crucial as the ones from the usability tests with the digital prototype. In general, the results from paper prototyping indicate that the users need to clearly understand what the application is meant for in order to agree to register to it with their personal phone number or insert any details of themselves in a user profile.

The user profile feature was only tested in the paper prototype and was not implemented in the digital version. Profiles in general appeared to be a disliked feature, as many applications ask for personal details without clearly stating what they are needed for. In SoAR, profiles could be used for sharing information about one's professional skills and finding experts. Implementing profiles requires careful thinking, as the industry users did not find them particularly attractive, either.

According to the paper prototype users, the application should provide visual clues, unobtrusive info texts or brief instructions during the first use to strengthen its image as a communication tool. Most of the paper and digital prototype users were suspicious about applications that request a phone number for logging in, but almost all of them agreed it felt more sensible when the application was identified as a video call tool, reminiscent of other voice or video communication apps. In general, users were more used to email authentication or the OAuth protocol, which allows authentication to third party services using for example Google, Twitter or Facebook accounts.

The users of the paper and the digital prototype alike considered taking and sharing snapshots of SoAR calls a useful function to be implemented in the future, and the opinion was later backed up by the industry users. With the paper prototype, we tested whether it would be feasible to store and share the snapshots from within the application instead of the native image gallery of the mobile device. However, users appeared to be at home using the native gallery, which is currently the common way to deal with images taken in mobile applications. It might also be useful to take photos and draw on them even when one is not having a call with someone.

On a more detailed level, there were four evident design changes that needed to be done to the SoAR prototype on the basis of the test results with the digital prototype.

*1) Controlling the visibility of the drawings.* All users were unanimous that automatically disappearing drawings were a serious disturbance. Increasing the erase timeout from 2 to 6 seconds did not make a crucial difference. As a result, we added a button for discarding all drawings from one's own screen: the hypothesis was that the drawings were quick drafts that could easily be reproduced should something useful get accidentally discarded.

*2) Stabilizing the drawings.* It was apparent from both the video recordings of the test situations and from the user interviews that the drawings needed to be stabilized in relation to the objects in the view. Only drawings that stick to the details they are meant to emphasize are useful. There were two options for world-stabilizing the drawings: mapping the drawings into the 3D space, so the users could move around objects and the drawings would stay put on objects, or pausing the video stream and allowing the users to work with details on the still view. Both options were voted for by the usability test users – however, as the industry users preferred pausing the stream, this was eventually implemented in the digital prototype. A user was able to pause her own camera view, which appeared paused on the conversation partner's screen, too.

*3) Improving the screen arrangement.* Users reported that the screen arrangement was not clear enough. Changing from one's fullscreen view to the conversation partner's view was clumsy because this had to be done by exiting fullscreen first. Also, it was surprisingly difficult to determine which camera view was one's own, since the views were the same size in two-view mode. We rearranged the view so that the conversation partner's view was fullscreen by default and one's own view could be seen in a small window on the lower right corner of the screen. Fullscreen views could be switched by tapping the small window.

*4) Introducing premade pointing tools.* Circling objects was the most common drawing action done by the users. As noted in most of the interviews, it is frustrating to repeat the same drawing over and over again manually Premade pointing symbols were considered useful. Helper 4 also mentioned that users might be put off by the crude quality of the freehand drawings – it is hard to draw exactly on a small mobile screen. As the users also experienced difficulty in describing directions, we added the option to use arrows instead of circles for quick pointing. It was possible to control the length and direction of the arrow by dragging a finger across the screen.

At this point we did not want to add features that would allow remote control of the conversation partner’s view. Only one’s own screen could be purged from drawings, switching screens affected one’s own view, each user could pause only the stream from their own camera, even if the paused frame was visible to the conversation partner as well. I wanted the new features to be as simple as possible in the field test phase and if the industry users identified the need for more control of the partner’s view, this could be implemented later.

# 5. SOAR IN THE WORKPLACE

This chapter is based on the field test findings: interviews and job shadowing observations are analyzed thematically, and SoAR calls and the perceived use of drawings in work context are reviewed in relation to the usability findings. The aim is to chart what the current communication practices in construction, facility maintenance and quality management are like and how SoAR could improve remote collaboration in these environments. The impact of the findings on further adoption possibilities and design choices is also assessed.

## 5.1 Existing Communication Practices

As mobile devices and specifically smartphones have become a standard tool in most lines of work, also the SoAR test users have experienced how mobile use has changed work practices. Combined with the development of web-based services this has meant an increase in the number of communication and work management systems, which can be used on multiple platforms.

The ISS maintenance person's experience of mobile use is a concrete example of how communication and work management are intertwined on the mobile platform. The maintenance person is able to operate several facility management systems on his phone. These systems are used to record and track the progress of each service call and the automatic notifications from them are a key source of daily tasks. Many facility automation systems also make notifications of their state to his phone. As a result, certain malfunctions can be taken care of before the users of the facility have even registered them. Even if web-based facility management systems reduce the number of phone calls from the clients, they do not completely replace live communication between the client and the maintenance person. Instead of being personally available for all clients, the maintenance person can now use his own judgment when giving out his mobile work number.

The surveyor at Skanska explains that a clear asset of modern mobile technology is the portability of up-to-date digital plans, which he is able to store on a tablet that can be brought along to the site. In general, the surveyors are more dependent on versatile use of mobile technology than other non-supervising employees.

Carry-along knowledge has a communicative effect, too: fellow workers rely to the surveyor frequently as a data source, because they know he has access to the latest plans. This communication pattern has existed before mobile technology, as well - surveyors have been considered to be able to provide other workers with technical details most effectively: “The foremen, it’ll take them hours if they go on searching some information. But the surveyor has it all in his head: measurements, images, structural engineering data, so he’s the one to ask from. Plumbers, electricians, they always ask the surveyor.” (Skanska shop steward.)

When it comes to applications that are specifically meant for communication, construction and maintenance workers used a fraction of the amount of apps that the Nokia quality managers needed on a regular basis: for example screen sharing, microblogging and video conferencing applications were only mentioned by the managers. The type of work affects the communication software needs – managing, supervising and organizing tasks require different tools than physical, mobile work. However, also the company culture plays a role. Nokia is a technology company, which shows in the attitudes of the employees: “People [here] find everything new interesting --- [they] are into trying out all these new possibilities.” (Quality manager.) On the other hand, on the company level Skanska has acknowledged the low degree of adoption of digital tools among the employees, and the company attempts to advance good technological practices through development projects. Even if most of the work conducted at construction sites is physical, I find it telling that the Skanska foremen do not utilize such a variety of communication apps as Nokia or ISS managers, although a major part of their work consists of managing, supervising and organizing.

The relatively low IT adoption rate is not only Skanska’s problem in the construction sector. When it comes to mobile technologies at construction sites, a key agent is occupational safety. The interviewed construction workers and apprentices did not report complete bans of mobile devices, but stressed safety procedures and meaningful, work-related use. At Skanska sites, the employees are instructed that, if calling is necessary, they ought to find a safe calling spot away from passages and machinery. Especially when working in the “pit”, in the foundation phase of a building, it may be dangerous to concentrate on a mobile screen instead of the surroundings. BauABC apprentices reported possible suspicion from older workers, should an apprentice be spotted with a mobile device in his hand at a site – nonetheless, they were optimistic that it is possible to convince the colleagues by explaining why mobile use is necessary.

Generation gap manifests in attitudes towards technology: “On the other hand we need to tell the older guys to use their mobiles more and with the younger ones, we should go and grab those away from their hands from time to time.” (Skanska shop steward.) Even if informal communication in social media is often seen as a distraction, it can also be harnessed to improve workplace communication culture. Informal communication at work has the potential to improve collaboration by building the sense of connectedness and social support, as well as generate new common ground (Zhao, Rosson 2009, p. 252). Some of the BauABC apprentices reported having successfully introduced instant messaging with the currently popular mobile app WhatsApp to their work community. Being able to communicate swiftly with a group of people, sharing images of possible problem areas and also using the app for funny, social content were assets that won over older and younger colleagues alike.

The Aalto lobby attendant has a similar experience with WhatsApp: she has used the app in her free time but has also started to promote its use at work after realizing instant messaging could make contacting the lobby staff more straightforward. In general, she feels the number of different apps is not an issue as long as they can all be used on the same platform, the smartphone. WhatsApp has also proven to be useful in maintenance work, especially for sharing photos in groups. Sharing visual content to build common ground has become frequent due to smartphones, in general. In all participating organizations mobile photos were used at least to some degree, but they were not always shared through an information system, instant messaging service or email: photos also served as a personal memory support. A common trait was using visual content to document faults and problems.

In urgent situations, phone calls remain the most common communication channel for all test users. Routine and ease of use when working in mobile, physical conditions are the key benefits of audio calls. Based on on-site observations, construction supervision and facility services are prime examples of fields where continuous phone conversations have a strong impact on the course and rhythm of the workday. For example, during the 4-hour job shadowing session with the ISS maintenance person I observed almost 20 incoming calls to his work phone, mostly from clients and ISS colleagues. Mobile phones allow constant availability, which carries the risk of exploitation. Knowing the employees can be reached at all times can result in less consistent work plans, and the supervisors end up giving small amounts of tasks over the phone instead of agreeing on a larger whole. At Skanska the issue is related to the relatively large number of own employees, who are expected to be more flexible than subcontractors. However, according to the foremen and workers alike, the low number of subcontractors compared to many other Finnish construction companies is a very positive thing and has more assets than flaws.

The high number of subcontractors and suppliers is a common characteristic in construction, maintenance and quality management alike. Supply chains may be long and complicated, and the most frequently cited risk factor in supply chain management at least in the construction industry is inadequate communication (Aloini et al. 2012, p. 744). Subcontractors may change rapidly in a project, and even if there was an effort to work with known and established subcontractors only, this may not be possible due to the resources of the respective companies. Each contracted company has a different communication culture, more or less compatible with the main contractor. Failure with communication can lead into tasks that are in the gray area, not clearly assigned to any partner in a project. Construction and facility maintenance workers also need to deal with several property owners and developers – contents of each contract with these top level companies may vary, which has an effect on not only the details of work but communication, as well.

When developing a remote collaboration tool like SoAR, it is essential to be aware of not only the communication patterns that could be reformed for the better by using the tool, but also those patterns that should not be mingled with. Face-to-face collaboration with colleagues or clients is typical in physical work settings, and two partially overlapping application areas are common to all participating organizations: problem solving and strengthening of interpersonal work relations. Meetings and other activities that are related to distribution of work are not taken into account as the focus is on physical work.

Face-to-face problem solving enables sharing the common view and effective gesturing, which results in higher situational awareness, like the construction surveyor describes: "Bigger issues, --- well, up until now it's like, I basically come to the office to have a look. So we can look at the images and --- [I] can explain it, because you can't do everything on the phone. By just talking, the other person can't understand what's going on." Nokia manufacturing facilities and their working conditions are inspected in co-located teams rather than on the basis of reports, and assisting an Aalto client with printing devices or clarifying specific faults in a facility is most effective side-by-side.

Meeting with colleagues or clients is also essential in order to maintain trust and to lower the threshold for future communication. Examples of this kind of behavior include the construction foreman's rounds at the site during which he talks to each and every subcontractor and Skanska worker to check how they are doing. The maintenance person's check up calls at the facilities serve a similar function. In a shopping mall maintained by ISS most service calls are addressed to a security person instead of the maintenance person, precisely due to presence: the security person is constantly available whereas the maintenance person visits occasionally and is not as well known to the shop assistants. The effect was also confirmed by the Aalto lobby attendant: after the owner of the building renegotiated the service contract due to tendering, the staff did not get to meet their new maintenance person for a considerable time, which had a negative impact on the client experience. Physical presence is an obvious part of the lobby attendant's work, as well, as it allows her to act as the reception, the help desk and the deterrent against potential crime attempts.

Personal presence is also a key element of system audits at Nokia, not only due to the effectiveness of unmediated information gathering. Introducing the audit team and the staff in charge at the facility is a part of the process, as well as discussions in person where the facility representatives can have a say to the findings immediately. Face-to-face communication builds trust and diminishes the effect of cultural differences compared to for example email discussions or reports.

All in all, SoAR can hypothetically support and even replace physical presence in many of the described problem-solving situations because it allows shared view and gesturing. However, it is less likely to substitute live communication that serves strengthening work relationships, which requires a personal connection based on co-location. All participants acknowledged augmented video conversation as a potentially powerful tool at work, and they could in addition come up with several ideas on how to use it in their own work. Everyone was familiar with at least one existing video call tool and many used Skype for personal communication.

## 5.2 Expectations

> Let's say there's some electrical fault --- so it goes like this. First we tell [maintenance]: ' Yeah, we've checked the fuses, the socket doesn't work. ' Maybe the next day, the maintenance guy comes over and goes: ' Yeah, that's right, doesn't work.' And he'll call an electrician, which will take another 2-3 days until he comes. That process could be faster with the video call… I could say: 'Look at this, here are the fuses and everything looks ok.' He'd see what I see and do. Because [otherwise] he'll have to come over, like maybe I didn't know where to look. --- But during a video call he could say: 'Hey, there's another fuse box there and there, go check that out.' Not that we'd take [the work] away from them, but it would serve the building better. If it's some quick fuse or so and it takes days from someone to come and figure that out, of course it's better if I can find the fault and fix it.
>
> Aalto lobby attendant

As noted earlier in Chapter 5.1, phone calls are the main method to respond to acute work situations, and mobile photos are commonly uses for documenting problems visually in most of the participating organizations. Hypothetically SoAR could substitute calls and photos in some help-seeking occasions as it combines the call functionality with shared visuals. All test participants could come up with potential cases in which SoAR could be the most effective way to communicate.

The test users from Aalto Campus Services and ISS were of the opinion that SoAR could reduce the time it takes to handle maintenance needs of the facility. Shared view and pointing by drawing could enhance identifying problems remotely. The maintenance person could possibly decide the appropriate solution over a SoAR call: whether the issue could be solved remotely there and then, whether he should schedule a visit or whether he needed to pass it on to a subcontractor like an electrician or a plumber. Communication between the lobby attendant and the maintenance person was decided to be the goal of the field test, whereas using SoAR for improving collaboration with subcontractors could be a beneficial test case in the future, should research on SoAR be continued (Figure 17). SoAR calls were aimed to supplement, not replace written records of service calls saved in the facility management system: these are necessary for following the overall service processes defined in the ISS service contracts.

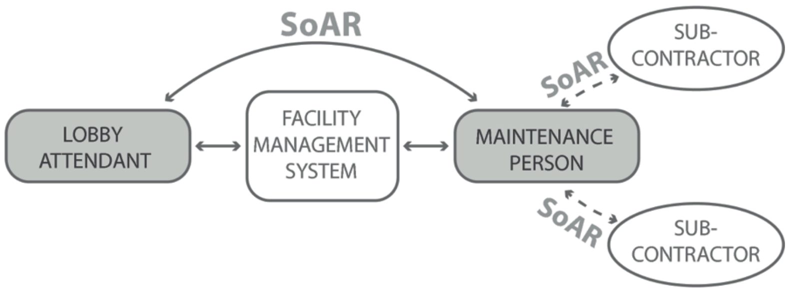



**Figure 17:** Aalto Campus Services and ISS facility maintenance test case

The Skanska workers speculated SoAR would be most useful during the frame phase of a construction project. The frame phase is critical for the stability and structural functionality of the finished building, because all the wall elements are installed at this time. Problems in this phase are prone to multiply should they not be addressed in time. The surveyor is typically the person who discovers possible inconsistencies between the physical site and the site plan and communicates these issues to the supervisors. The frame phase requires exact scheduling and good collaborative readiness from all contractors (Lehtinen 2009).

The Skanska participants suggested that, if discrepancies are found the plans, SoAR could enhance communication between the structural engineering agency that prepares all the plan documents and the site representatives, not only between the on-site personnel. The foremen are currently responsible for contacting the agency, and a typical bottleneck is the time it takes to update the plans. If the agency cannot act quickly, the main contractor may need to come up with a workaround to be able to continue working despite the faulty plan. In such cases the realized structures are updated in the plans afterwards. With a tool like SoAR a surveyor could possibly communicate directly with the engineers to collectively come up with the best solution in a conflict like this, which could lead into faster and more sustainable solutions.

Long distances could also be mitigated with SoAR: at Skanska Härmälänranta site the site office was located 50 meters from the actual site, which was considered unpractical. Because the foremen did much of their reporting and coordinating tasks on computers at the office, situations in which remote collaboration between workers on the site and the foremen were expected to occur on a regular basis.

The aim of Skanska test case was, that the surveyor and foreman would use SoAR to solve issues noted by the worker teams on the site, and direct communication between the surveyor and the structural engineering agency could be a possible test case in the future (Figure 18).

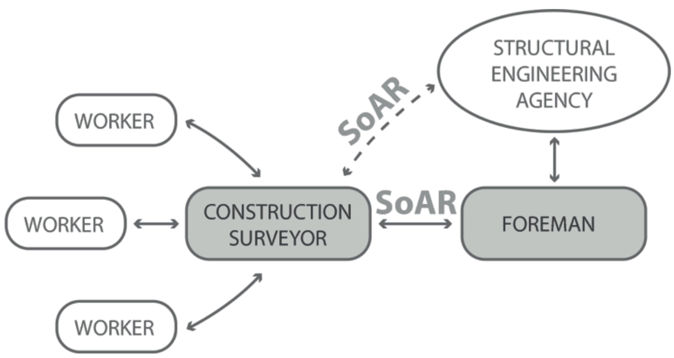



**Figure 18:** Skanska test case

The initial idea for testing SoAR at BauABC was to find out whether the application could be used for extending the field of view of construction machine operators. Due to blind spots, BauABC apprentices needed to give directions from outside the cabin to the operator, for example when he was maneuvering the shovel of an excavator or installing sheet piling into to the soil in the right angle. The instructions were given with walkie-talkies and pointing gestures. SoAR could possibly replace the walkie-talkies: an apprentice could stream the outside view and instructional drawings to a mobile device mounted inside the cabin.

At BauABC we only had half a day to test SoAR, so I chose to set up a fairly straightforward help-seeking case instead of testing whether SoAR could be used as a vision aid of the machine operators. The BauABC field test was the only case that took place in an educational setting, so I wanted to find out whether SoAR would enhance reporting progression and communicating possible problems like equipment malfunctions to the instructor in a distributed task setup (Figure 19).

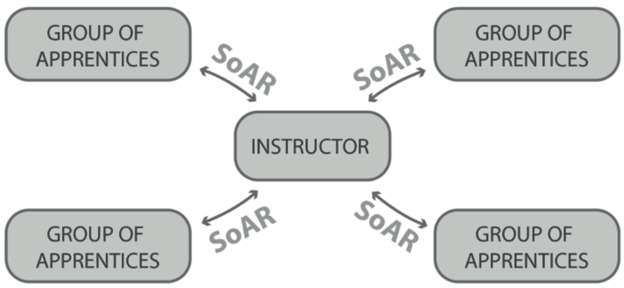



**Figure 19:** BauABC test case

The quality managers at Nokia expected SoAR to be useful when auditing for example Nokia's own production lines, site installation services, maintenance services or logistic units. Audits at external, contracted supply units that are not part of the Nokia corporation are more challenging, because taking video and pictures is often strictly regulated. Even if SoAR did not record video by default, trust issues could build up.

Currently, the auditing teams in internal audits use mostly written notes to report their findings to the audit lead. With SoAR it would be possible to collaborate remotely to distinguish actual, relevant on-site findings from secondary issues immediately during the facility round, instead of relying on notes afterwards. This would potentially speed up the audit process and thus improve its coverage, because according to the managers, it is common to run low of time during these inspections. SoAR could also improve efficiency by making better use of distributed knowledge.

Auditor teams may suspect a finding in an area they do not have expertise on, but can confirm the finding by making SoAR calls to the audit lead or a colleague specialized on the topic. Also the local facility managers could be involved in responding to the suspected findings early on. In some cases even lightweight audits done with SoAR could come into question: a local team could present the facility to a remote auditor.

SoAR could also be beneficial when following up on the required improvements after the audit. Especially in cases where the audit follow up is a written report and not a new visit, reviewing the facility remotely over a SoAR call could provide the auditors with additional affirmation. The quality managers anticipated SoAR to be useful for bridging cultural differences, as well - shared view would allow building common ground quickly, whereas communicating problems and development issues in writing can be ambiguous between representatives of different cultures: “There are still massive differences in how people act in different places. Even if we basically have common processes and rules, the environment has an effect. So visual communication is a ‘great equalizer’ of sorts, in that sense.”

The Nokia test case was decided to consist of simulated audits, which would aim to verify whether SoAR could improve collaboration between auditing teams and the audit lead (Figure 20).

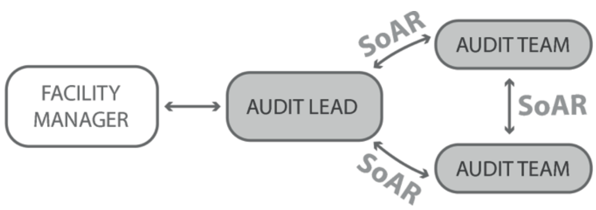



**Figure 20:** Nokia test case

Saving time and money by for example reducing the need for immediate presence was a common expectation for SoAR. Although SoAR hypothetically has the potential to make work-related problem-solving faster, the emphasis is rather on improving workplace communication than enabling savings for the employer. The study concentrated on the employees’ point of view: in case the employees considered saving time with SoAR essential for making their work more manageable, time was considered an important aspect.

## 5.3 SoAR Test Calls

### *5.3.1 How SoAR Was Used*

The industry users made altogether 18 SoAR calls from actual work conditions during the field tests. In addition, they made facilitated test calls with each other before the field test period and some check up calls with me during the test phase. The actual calls made during the field tests were mostly simulated, made for the purpose of making a SoAR call and not necessarily to solve an actual problem. The calls had work-related content, though, as the actual work environment was the physical framework in every call.

The calls could not be analyzed to such detail as the usability test calls because of technical issues related to the recording process of the calls. There was a recording malfunction during the first field test period at Skanska, due to which almost all video and audio files were corrupted. All users were instructed to turn off the recording in case it deteriorated the quality of the calls too much – this resulted in partial documentation of the calls made in the later test phases. I had enough recordings to understand the context of most of the calls, though. As the limitations of the recordings already came up during the first field test phase, I could compensate by paying more attention to the final interviews after the test periods. Majority of the field test results are thus based on discussions on the user experience and the future prospects of SoAR.

Skanska test users were not able to verify the hypothesis that SoAR would be useful for improving communication in the frame phase, because the surveyor and the foreman were working on different sites at the time the test period started. The surveyor moved on to stake out the foundations of a new building while the foreman continued to supervise the interior phase of the apartment block site they had previously worked together on. They made four SoAR calls to each other, during which they tested how the application worked inside the building, out in the pit and in different lighting and noise conditions (Figure 21).

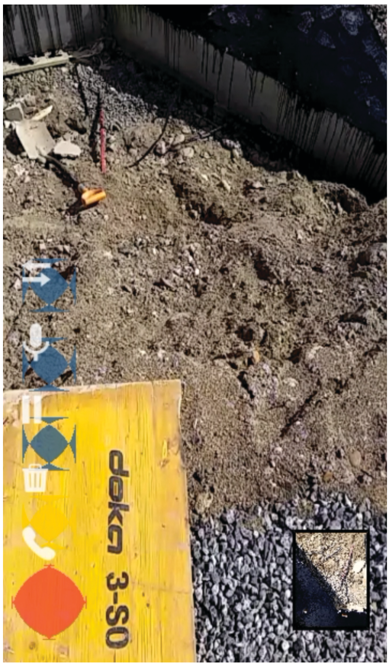


**Figure 21:** Screenshot from a call at Skanska. The icons are malformed due to the recording feature: the actual UI of the users can be seen in Fig. 7, Chapter 3.5.

The test users of Aalto Campus Services and ISS facility maintenance agreed to address the maintenance needs of the Miestentie facility with SoAR calls. The lobby attendant was highly motivated to test SoAR and came up with several topics for calls. Scheduling was especially challenging in this test case: the users were occasionally working different shifts, which limited the time frame they had available for calls. In practice it was necessary to schedule each SoAR call beforehand, because the maintenance person had such a tight timetable during his workday. Despite scheduling attempts, the participants experienced difficulties in making as many SoAR calls as they had planned.

The lobby attendant and the maintenance person made three SoAR calls during which they tested the basic functions of the app. Two calls handled actual maintenance issues (Figure 22). The lobby attendant reported burnt out exit lamps, which had been noted during the fire inspection done in the building. In the other call the attendant showed the water tap outside the building and requested an adapter for attaching a hose to it.

The issues that were handled with SoAR were fairly small and it was not yet possible to prove that using SoAR speeded up the maintenance process.

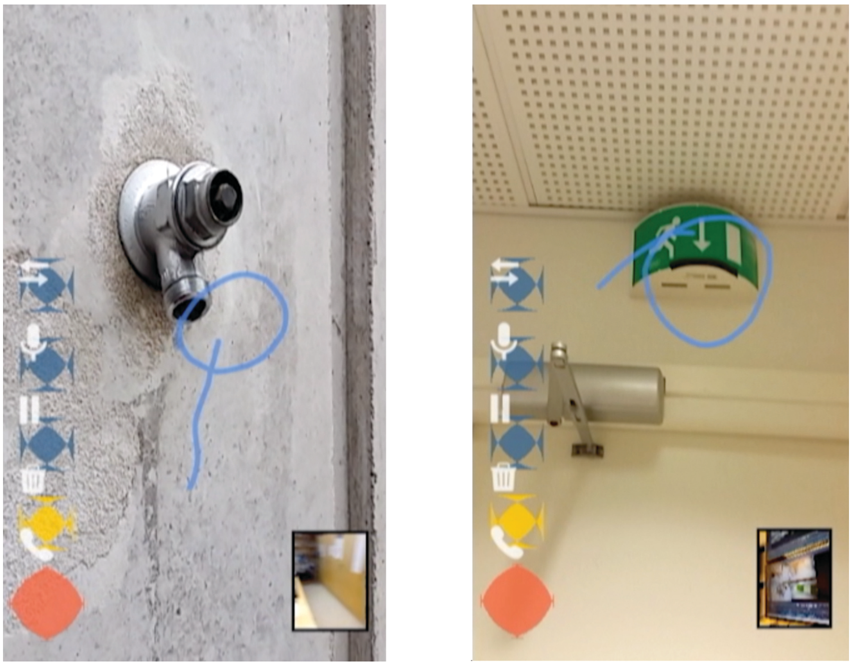


**Figure 22:** SoAR screenshots of Miestentie maintenance issues

The quality managers at Nokia decided to present SoAR to their colleagues to develop use case ideas further. There were two test cases in which they used SoAR: the first was a workshop held in Germany, in which the idea was to test whether SoAR could be used for including a remote expert in the workshop, and the second was a simulated audit in the Nokia premises (Figure 23). Viewing documents and diagrams was more essential in the Nokia use cases than the ones in the construction and maintenance sectors. The quality managers tested viewing both printed documents and those on a computer screen. Interestingly, the call from Germany to Finland worked technically very well, but the calls within the Nokia building in Espoo often suffered from missing video or audio stream.

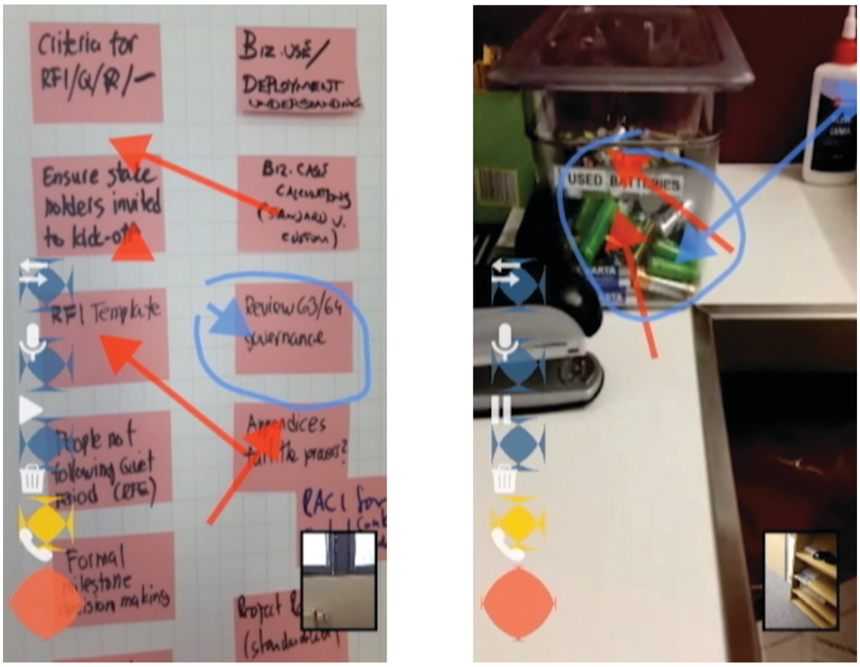
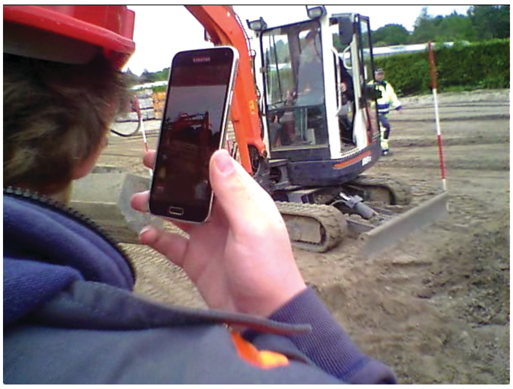



**Figure 23:** Workshop notes and a simulated audit tour at Nokia (screenshots from recorded calls)

The instructor from BauABC was committed to testing and came up independently with different uses of SoAR. He tested drawing during each call in a meaningful way by pointing out details and also by asking the apprentices to show details of their work by drawing. The whole field test at BauABC was facilitated, which proved to be an effective choice for collecting as much material as possible over a short period of time – the call material from this test case was eventually the richest, even if it was not possible to conduct as thorough background and feedback sessions with the apprentices as with the other test users. It was also possible to avoid most of the technical problems experienced in the other test cases. I prompted the apprentices to make seven SoAR calls to their instructor from the task locations (Table 7).

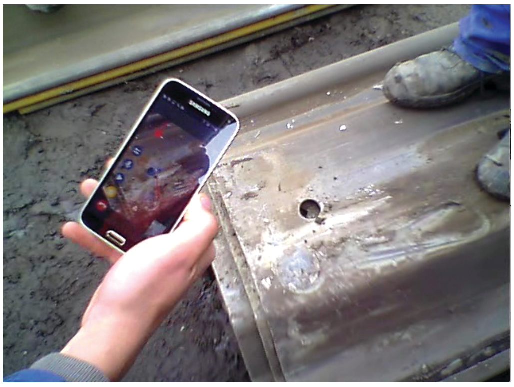
**Table 7:** SoAR calls made at BauABC, Rostrup (video screenshots)



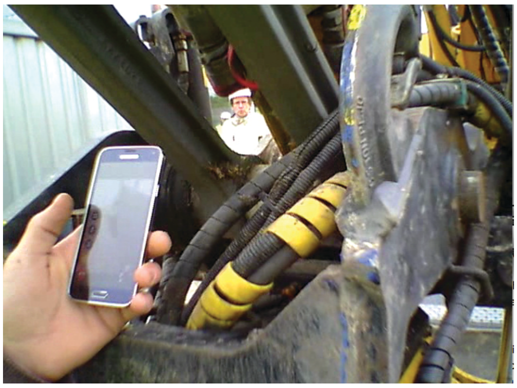
**Call 1**
"We're currently at the mini excavator and making the 10 meter trench and we have a small question about it, namely the joint up there. We oiled it yesterday but it squeals like a pig." The group and the instructor figure out together the cause of the noise by looking at the joints and pointing at them by drawing. The instructor asks questions about the procedure the apprentices have followed and whether they had lubricated all joints. The apprentice in the photo struggles to hear through noise from the machines.

**Call 2**
The previous group shows the mini excavator again. The instructor gets to see the joints and cylinders closer, and he finds out the cylinder is the cause of the noise. Lubricant gets pressed out unevenly: the group looks at the underside of the joint, to notice the right side is not properly covered with the lubricant. The instruction is to lubricate the parts properly again.

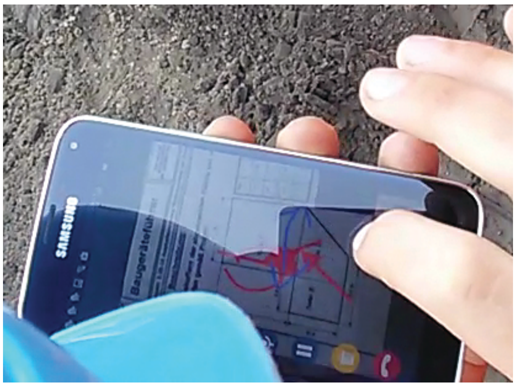
**Call 3**
"We have a small problem, too little oil in the excavator. We need more oil after the pause." The instructor asks the apprentice to show the problem, the amount of oil in the engine. The apprentice checks the oil while the instructor asks questions about the kind of oil required for this specific machine.



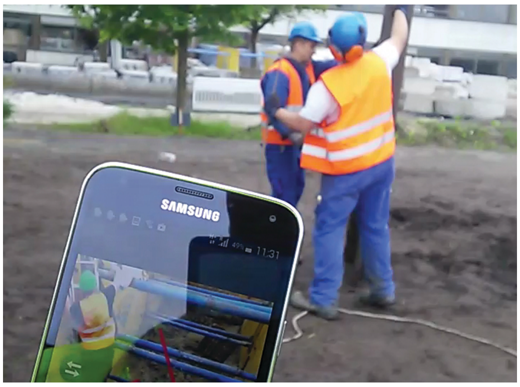
**Call 4**
Situation check with the sheet-piling group, which has just begun the exercise and does not have issues or problems as of yet. The apprentice shows the sheets to the instructor, who draws and asks about the marks on the sheets - these were caused by the grip of the vibrating unit.



**Call 5**
Screws are missing from the excavator - apprentices do not think this is serious but decide to show the situation to the instruction over SoAR.
There is a malfunction: the instructor only sees black on his screen, even if the apprentices see the stream from his camera.




**Call 6**
Situation check at the excavator. The instructor rides with his bike to the location where he has the written instructions with graphs of the tasks. He goes through the set of instructions with the apprentices, and both draw to find out how much of the task the apprentices have already accomplished.




**Call 7**
Situation check with the sheet-piling group: in the middle of the call the apprentice is needed to help adjust the position of the sheet. The SoAR phone ends up to me, and the instructor explains a bit what is going on at a different excavation, with drawings.

Despite technical issues, SoAR was considered an easy application to use by all field test participants. Even if the technical problems reduced the amount of data from the actual calls, it was essential to learn how the prototype worked in long-term independent use before planning any wider scale test cases. The BauABC case was a strong indicator that facilitated test situations were most productive at this stage of the design process. The calls made at BauABC also showed the potential of SoAR in actual problem-solving situations: according to the apprentices, troubleshooting the malfunctioning cylinder of the mini excavator worked well with SoAR as they could locate the problem and service the cylinder with the remote help from the instructor.

### *5.3.2 Drawing on Video*

The SoAR calls made during the field tests were generally not as task oriented as the calls made in the usability tests. This has an impact on how the participants used the application, but due to the more informal nature of the field test calls and the small amount of recorded material, it is difficult to make detailed conclusions on how the users utilized the drawing feature in relation to verbal communication.

What is clear, though, is that the primary purpose of drawing on screen was pointing out objects and locations, which is in line with the findings from the usability tests. In all recorded calls pointing gestures were in a predominant role, and the users also expressed the preference for pointing over freehand drawings in the interviews. Premade arrow function was found practical and easy to use. Circling was also a very common drawing gesture among all users.

During the recorded calls both conversation partners typically utilized drawing, as a contrast to the usability tests where the task setup reduced drawing only useful for the remote instructor, the helper. Obviously the field test users were also keen to test a novel feature, which lead both conversation partners to draw. In the usability tests the so-called workers who conducted the actual physical tasks systematically preferred pointing at task objects and locations in the real world, not by drawing on screen. Such a clear division could not be discerned in work-related SoAR calls. I assume that when an object is large or out of reach, like an exit lamp presented by the lobby attendant, or when it is important to transmit as broad a view of the physical surroundings as possible, it becomes more practical to point out details by drawing than by pointing at the actual, co-located details themselves. The usability test users worked with fairly small objects only: Lego bricks and magnetic characters were tiny and the surroundings of the task items were irrelevant for solving the tasks. It was thus convenient for the workers to operate close enough to be able to grab the objects and point at them directly.

We made the decision to stabilize the drawings in SoAR by pausing the video stream instead of making 3D stabilized drawings. All field test users tested pausing the stream and found it to be an essential feature for successful drawing with SoAR. The users also found an important benefit in the way that the pause function works. Freezing the view not only mitigates shaking and helps to draw in more detail; it also allows the user to communicate details that are physically hard to reach. Instead of staying in an uncomfortable or even harmful position to continuously stream the view, the user can freeze the details on the screen, move to a better spot and continue to discuss and draw on the still view. The LED light of the mobile device could be used for lighting dark spaces during the calls.

## 5.4 Augmented Video Calls at the Workplace: Advantages and Challenges

### *5.4.1 Construction Sector*

Even if all the original use case ideas in construction work were not tested, the results from the field tests suggest that SoAR is both usable and useful in this line of work.

The BauABC apprentices were still interested in testing SoAR further in the future, as a vision aid when operating construction machines, and the Skanska users were positive after the testing that communication during the frame phase would be improved with SoAR, if the application was adopted in the work community.

SoAR worked especially well in construction education settings where the exercises were conducted independently in groups distributed over a large area. The apprentices could contact the instructor effectively and he could assist in actual technical problems remotely due to the shared view and drawing option. The instructor was also able to monitor task progress remotely. The apprentices also believed the application would help them in their actual work, for example when encountering technical faults in the machinery and communicating these to the personnel at the repair shop. Machine operators also need to contact engineers on a regular basis, e.g. when encountering geological problems during excavation work.

Although the Skanska workers concentrated mainly on testing how SoAR fared in the actual construction environment, instead of solving actual issues related to the site, it was evident that new application areas for SoAR would be uncovered after extended use. For example, Skanska has well-established practices for monitoring occupational safety, and safety observations could be effectively shared with SoAR. Instructing new, sometimes inexperienced subcontractors requires a considerable amount of time from the foremen. While SoAR cannot replace face-to-face communication completely, it allows more effective assistance and status monitoring in such situations. One of the assets video stream has over more familiar still images is the chance to view structures from all angles, by going around them with a camera.

Harsh weather conditions are a challenge in construction in general. Earlier research findings on environmental challenges for mobile use in construction work were confirmed in the field tests (Pejoska et al. 2016, Bauters et al. 2014). Building noise was the most frequently noted issue that complicated using SoAR. Machinery makes a lot of noise and wearing hearing protection, while compulsory, hampers mobile use. The apprentices at BauABC had a positive attitude towards solving the noise issue  - they considered wearing plug headphones under the earmuffs a viable solution. Earmuffs with embedded headphones and microphones were currently not common at Skanska or BauABC and could not be tested together with SoAR.

Direct sunlight makes seeing the video screen difficult, but this issue cannot be alleviated much with current mobile screen technology. Rain and low temperatures were also discussed with the Skanska users - mobile devices can handle light rain, but water drops tend to interfere with the capacitive screen and how it registers touch. The mobile devices currently used at Skanska function well in all temperatures – these can vary between -30 and +30 °C at Finnish construction sites. In physical work context there may be situations where having only one hand free and one holding the phone will become cumbersome. Protective gloves are also a hindrance, because they need to be taken off to answer a call and to draw. These challenges are not specific to SoAR use, as they will be encountered with most mobile-based communication methods.

Areas with large construction sites like Skanska's do not necessarily have network infrastructure in place until the buildings are in their intended use. Even if the mobile network were sufficient outdoors, the connection may be unreliable in the indoor locations of a construction site. Connectivity problems were also familiar to the BauABC apprentices, although these were seen to be more of a concern in rural areas of Germany. Mobile network coverage was sufficient in the construction sector field tests, which was encouraging for possible wider adoption.

All in all, mobile devices are readily available and work sufficiently well in the demanding conditions of construction sites. Although mobile devices have their limitations, they are the robustest technology currently available for shared-view communication in the construction sector.

### *5.4.2 Facility Maintenance*

The ISS maintenance person and the lobby attendant from Aalto Campus Services experienced SoAR useful, despite technical obstacles during the field test. The aimed use of the app was reporting service needs of the Miestentie facility. Because each service call made from a facility needs to be filed in the facility management system as well, using SoAR lead to communicating service calls twice. Although the lobby attendant considered these double announcements a downside, this is actually not in conflict with the current communication practices. The maintenance person and the lobby attendant communicate the service needs also by phone on a regular basis, and SoAR would thus have potential as a more effective replacement for the audio calls.

The users came up with new use case ideas after the field test period. Part of the lobby attendant's work is instructing teaching personnel how to use the presentation equipment in the classrooms of the building. In case she cannot go and assist the teachers personally, a remote tool like SoAR would be powerful. Especially when helping remotely with switches, sliders, cables and RC units, the lobby attendant considered pointing out details by drawing a valuable feature. Her initial expectation that SoAR could be used in acute situations like reporting a failure with the ventilation system was strengthened. Recognizing the actual part of the system that has a malfunction with SoAR, before the maintenance person is able to arrive, can prevent more extensive damage.

In facility maintenance, an especially good future case for SoAR would be handling the service calls of a new building. The guarantee period of a new construction in Finland is two years, during which the main construction contractor addresses most faults in the facility, even if the faults are reported through the facility maintenance company. As the contractor is already familiar with the facility due to the construction period, many issues could likely be effectively identified and furthered over augmented video calls between the maintenance personnel and the contractor representatives.

The maintenance person stated that communication with established professional contacts like subcontractors and property management agencies is often more fluent than communicating directly with the users of facilities. The latter may be very motivated to collaborate and participate in keeping a facility in good condition – however, they do not necessarily have incentives to do so if their work is not directly related to looking after the facility. The maintenance person believes he would benefit from having a tool like SoAR to communicate with motivated users of facilities, but a more realistic prospect lies in internal communication and managing the supply chain:

> I believe we would have use [for SoAR]. Between us and the supervisor, for example. Or us and a subcontractor, because it would reduce the work by an hour or two, from when someone comes over, checks it out, and goes out to buy it. If we can show it in a video call and that's enough --- Let's say we need a filter - and we don't necessarily mention the model or trademark. So you can't find it. But if I take a photo, it's possible to get it from the store

Interestingly, the snapshot feature that has not yet been implemented in SoAR during the field test is already visualized into the workflow described above.

### *5.4.3 Quality Management*

The Nokia quality management case provides a fruitful perspective that extends the use cases of SoAR from conducting and supervising physical work to monitoring and developing the whole supply chain, where both administrative tasks and work in physical environments like production facilities play a role.

After the field test, the quality managers felt that SoAR would be optimal for confirming findings from an audit to a supply facility – not necessarily during the actual audit rounds but at the summary sessions after each of the audit workdays. Due to time restrictions the live rounds cannot depend on a technology that might not function reliably. Even if the application itself worked flawlessly, the mobile network might not be sufficient between two countries due to limited bandwidth and roaming access.

Video stream was seen to have important assets compared to still images. When reviewing the required improvements after the audit, images can be unreliable as proof. Reviewing the improvements through video stream was considered more revealing, as the auditor can request the conversation partner to show different locations inside the facility. The drawing feature allows marking the points of interest.

SoAR was also introduced in a workshop setting, where remote participants would typically join in through a video conferencing application. SoAR was tested outside the workshop agenda, so it was not possible to learn about the remote participant's sense of telepresence in the course of the workshop. The drawing feature was found out to be useful for discussing details of the workshop notes, although using SoAR for viewing detailed documents was problematic. The screen of a mobile phone is quite small and the automatic focusing of the camera lens can cause a frantic effect on fine details like letters. Paper documents were easier to view than screen documents, which tended to flicker. SoAR fared quite well when compared to using a laptop camera, though: it was easier to zoom in with a handheld device.

Viewing documents like instruction sheets, warning signs and quality standards is important in auditing work, but the quality managers were confident that this part could also be covered with other methods than SoAR. SoAR was considered a promising tool for making observations on physical work conditions at for example production lines, but it was not expected to cover all communication needs related to audits.

The Nokia users believed, like other field test participants, that using the app further will generate more use cases. SoAR is seen as an optimal tool for impromptu conversations, sudden situations that require a remote expert. Possible additional uses include solving logistics problems or acute situations at production lines and handling delivery complaints in case of broken products. Remote problem solving is also often needed on hardware installation gigs and in the research and development laboratory environment. These cases are interesting for future research on SoAR: shared view and pointing by drawing could likely improve remote collaboration in work contexts abound with technical details.

### *5.4.4 Adoption Challenges*

In addition to context-specific challenges, there are also more generic issues related to the wider adoption of SoAR. Training the staff to use certain digital tools in their work is a considerable investment for large companies like Nokia, Skanska and ISS. Due to this, such tools need to be not only functional but also fully productized to signal long-term reliability. Although it may also be possible to increase the use of tools like SoAR from grass root up, for example the Skanska participants believe that extensive official training is required before an application like SoAR can be effectively used.

The price of data streaming can be a concern especially in international companies like Nokia. Employees do not typically have unlimited data plans and even if they did, roaming will increase costs when working abroad. Unlimited data plans are currently common in Finland, but this is not the case in every country.

Windows was the official work phone brand at Aalto Campus Services, Skanska, ISS and Nokia, but due to limited development resources SoAR will not be available for Windows phones in the near future. Depending on the direction of the work phone market and how work with SoAR will be continued, it may be necessary to consider developing for Windows phones, especially if field tests in the industry are continued on a larger scale.

Studying the adoption of SoAR on an organizational level is out of the scope of this thesis. However, the field tests provided a hint of how fragile the use of a new application is. Testing gave a plausible view into what kind of technical and work-related obstacles there would be to using a more developed, widely available application, as well. As long as the use of an application is not routine, any setback can result in lower motivation to return to the application again – be the setback a sick leave, flat battery or a busy day during which testing does not feel comfortable. The Nokia users also believe people have grown more demanding when it comes to mobile technologies: these are dropped quickly if the first experiences are not good. All in all, looking into the wider adoption of an application requires understanding of the complex effects of the users' hedonic motivations, habits and experiences of price value (Venkatesh et al. 2012, p. 174).

## 5.5 Discussion

The aim of the field tests was to find out whether augmented video calls are useful for problem solving and communication in physical work context. According to the observations and feedback from the tests and interviews, the app was found out to be a potential solution especially for acute and ad hoc work situations in all the test environments. It can also be beneficial for improving communication in work processes like quality management and managing supply chains in the construction and facility maintenance sectors. Although there were challenges to using SoAR as well, these had more to do with the facilitating conditions and the technical state of the prototype rather than the application and its ideal use.

Analyzing the conversational characteristics of SoAR use in work context was not possible due to the limitations of the call data. Comparing the conversation analysis results from the usability tests with the field test data would have revealed whether the verbal and gestural patterns differ from each other in controlled and work-related contexts. Facilitating the test situations and deriving structured, controlled test tasks out of actual work situations in cooperation with the test users would have resulted in more comparable data. However, the field tests, organized as they were, produced a lot of information on potential use cases and the real-life conditions in which communication apps like SoAR are used. A more controlled test setup would have likely compromised much of this information as many of the challenges and possibilities came up due to the independent coordination of the testing conducted by the industry users.

Trying SoAR out leads into new potential use cases: “Now that I got to experience it, I’ve noticed quickly those situations where, again, it would have been useful to have.” (Aalto lobby attendant.) Other users expressed similar thoughts, as well: the reporting process of the faults of a new building to the main constructor during the guarantee period could be more fluent in the maintenance sector, construction surveyors could start to communicate directly and effectively with structural engineers, and verification of audit findings could be more conclusive with augmented video calls in quality management. In addition to being useful as such, SoAR and similar tools could also improve communication practices in the workplace in the future.

## 5.6 Design Implications

The most essential design input from the field tests deals with aspects that maximize situational awareness and common ground in remote collaboration taking place in physical work contexts: pointing by drawing and sharing the live view. These could be enhanced by making some improvements to the design.

In the SoAR prototype that was used in the field tests, callers initially saw each other's video stream. The first moments of a SoAR call can enhance situation awareness between the conversation partners even better, if the initial view on both users' mobile screen covers the acute work context that is supposed to be handled during the call. If the topic of the call is seeking help from a colleague, the caller typically wants to show something from her own surroundings to the person she calls. The users from BauABC and the facility maintenance case confirmed this assumption: the callers found it annoying they had to switch the view on the phone each time they made a new call. As the Aalto lobby attendant states: "When I called him, his stream was shown large on my screen. I was like, I don't want to see that, where's my icon. I always changed into my view." Having to repeat the view switch over and over quickly became a nuisance. However, there are other possible use cases for SoAR than just seeking help. In auditing, the quality manager might want to contact a supplier representative and see the view from the manufacturing site, instead of his own surroundings. SoAR should allow the users to configure which video stream they want to see when calling. However, the help seeking case - showing the caller view first to both conversation partners - should be the default.

On the basis of observations and interviews, the users often needed to communicate verbally which screen they were looking at or drawing on. The problem is partially related to the simulated nature of the test cases: an actual task or work-related problem would direct the activities more clearly. Still, marking the screen the conversation partner is working on is a small development effort that improves collaboration and reduces the need to explain actions that are not related to the task at hand and should be thus implemented.

According to Licoppe and Morel in their study on informal video conversations (2012), participants in video calls strongly expect to see the conversation partner's face in the initial setting, and the "talking heads" arrangement is also expected to reoccur during the call. The results from the SoAR field tests contrast with this notion: none of the users found the face-to-face formation mandatory for successful communication. Only the Nokia quality managers stated that the possibility to flip the face camera on would be beneficial in some occasions, which is understandable considering their use cases included e.g. a workshop. Facial view would also increase the sense of personal connection and trust in audit-related remote reviews. However, also the quality managers acknowledged showing one's face is quite personal and should in some cases be left to one's own consideration. Most of the industry users did not find facial view useful at all but preferred directing their attention to the physical work context: "I don't know why there should be [frontal camera view] --- I think it was ok like this, because it wasn't meant for having conversations like that, but it's about what the other person is doing" (Aalto lobby attendant). The users from the construction sector even found the thought of viewing a colleague's face amusing, which I believe reflects strong task orientation at work: "I'm really not into seeing what the bags under his eyes look like, whether he's had a good night's sleep or not" (Skanska foreman).

On the basis of the results, the possibility to flip the camera view is not a high design priority. Even if it was implemented later on, the facial view should not be the default view for SoAR conversations, but the attention of the users should be directed to the surroundings instead.

The users were generally satisfied with how the drawing feature was implemented. Premade arrows, freehand drawing and pausing the screen for stabilization were considered simple and sufficient, and only minor improvements were suggested. As pointing by circling was the most common freehand drawing gesture, the quality managers requested a premade, scalable circle in addition to the arrow feature. However, simplicity of the user interface is a priority when designing new features, which the Nokia users were ready to admit as well. In all the cases users reported having intuitively tried to zoom into a paused frame. This indicates that a zoom function would be very useful, if the resolution of the paused video frame allowed sufficient level of detail.

The immediate nature of video calls makes them effective for problem solving, even if still images remain important for documentation and later reviewing. The possibility to store screenshots with drawings was unanimously deemed necessary, and this feature was implemented soon after the field test phase. The images are stored to the default gallery of the mobile device, because the users stated sharing from the device gallery to other applications and platforms was intuitive and commonplace.

Design issues that were related to general usability in physical work context were also uncovered. The most severe issue was network and server connectivity during the field test period. Breaking calls and faulty streaming were common, as well as login problems. As noted by all users, improving the service quality in this sense is inevitable, should SoAR be used in actual work situations.

When SoAR was on, the mobile devices tended to heat up quite a lot, which indicates high power consumption. The users also reported the batteries of the test devices seemed to run out faster than usual. For actual work use, the power consumption of the app should be optimized.

A generic review of the user interface of SoAR is a necessary step in the future. The improvement needs described above have an effect on the interface, and also the locations and shapes of the current in-call functions need to be reconsidered. The users found out it was too easy to press the icons accidentally when on the move or when concentrating on the surroundings. Hiding automatically for example the icons that are related to managing the video call would allow more space for the video stream on small mobile screens.

# 6. CONCLUSIONS

## 6.1 Design Outcomes

The research-based design process of SoAR produced results on what features are essential for the users of video-based, augmented communication tools in work context and how these features should be implemented.

For immediate adoption in physical work context, an AR communication application should rely on currently available, off-the-shelf platforms like mobile phones and tablets. Emerging AR hardware is not yet mature enough for wide-scale communication needs and environmental challenges that physical work situations pose to technology. AR-based communication cannot rely on content that needs to be specifically prepared and updated for the application only. Thus, a lightweight augmentation like drawing on screen is a sustainable way to enhance visual communication in constantly changing, unprepared work environments.

Any application should clearly communicate what it is meant for, what data it will need and what the data is used for. A video communication app like SoAR can make good use of the mental model of a regular phone call app by presenting the user with contact lists and call logs. The feel of a communication tool makes it more acceptable for a user for example to approve the use of her phone number for registration.

SoAR test users deemed implementing user profiles as pointless, even if these could later on be used for finding experts in the work community. Sharing more personal information than absolutely needed was considered negative. Also this is due to the mental model of the application – should expert seeking become an essential function of SoAR in the future, the application must be developed so that the users understand and accept the concept of sharing expertise as the key feature of the app. At the moment this is not the case, as the current SoAR prototype has been developed with only the communication purpose in mind, to allow testing how augmented video conversations function in work context.

Maximizing situation awareness and building common ground have been the key principles for designing the digital SoAR prototype. Shared live view is the most important component in this pursue. Determining what the live view should be, though, is not straightforward: “talking heads” formation like in conference calls, each user’s own camera view like in camera applications or something else? The best option is, that both users see the same view immediately, the environmental view from the caller who is the person most likely seeking help. The application should also indicate the users which camera view their conversation partner is working on. Building common ground must be possible also after the calls. Thus, it must be possible to store screenshots of the video calls for later viewing.

Drawing on the touch screen allows the SoAR users to point out objects and locations, demonstrate directions and proportions and to come up with other clarifying figures to enhance situation awareness. The users must have control over clearing the drawings and stabilizing them in the video view. Pointing at objects and locations is the most common way to use the drawing feature, and pointing should be made as easy as possible by providing the users with premade pointing symbols like arrows or circles. A freehand drawing option is needed to cover more complex gesturing needs. Drawing and speech complement each other: drawings allow economical use of language but do not replace the need for verbal communication.

In general, SoAR has been designed to allow the users communicate their work-related issues as effectively and simply as possible. The test users did not need tutorials or lengthy introductions to start using SoAR. According to the findings of this thesis, a moderate amount of well-designed features results in a smooth user experience, which is essential for both usability and wider adoption possibilities of the app.

Drawings over live video conversations should be implemented more widely in video conversation applications in general, because they have strong potential in improving remote collaboration in the light of this study.

## 6.2 SoAR as a Workplace Communication Tool

The SoAR prototype was developed in an iterative, research-based design process, while working closely with users from construction, facility maintenance and quality management sectors. The application is generic enough to be used also in other lines of physical work: the users did not feel the need to adapt SoAR to their tasks. All design choices have been validated in a user-centered fashion and the feedback and observations of the usefulness of augmented video calls in work context have backed up the main hypothesis of this thesis: SoAR is useful for help-seeking and problem solving in physical work contexts.

Drawing on video stream was found both effective and preferable compared to communicating over video and audio by the usability test users, and these findings were supported by the field test results. The users believed pointing by drawing reduces the number of misunderstandings and builds common ground quicker than when relying only on shared view. However, having a video call application even without augmentations for viewing the work surroundings was found very useful. In the light of this remark find it interesting that none of the field test participants had used or knew examples of utilizing video conversations in physical work. In the future it seems worth digging into the reasons why video call applications are still seen as mainly a conference tool, not an aid in physical work even if the advantage of shared view was recognized.

The importance of the mental model of an app was described in Chapter 6.1. In all participating organizations, phone calls were the most important way to communicate in acute situations. As the mental model of SoAR resembles phone call apps and the purpose of SoAR is enhancing remote collaboration in immediate situations, the scenario of SoAR as an effective substitute of audio calls seems realistic. Using SoAR and other AR applications on current mobile devices would act as a necessary adjustment phase before moving on to completely different communication interfaces and mental models – wearable AR hardware like HMDs.

Although SoAR could theoretically replace audio calls, it cannot explicitly replace face-to-face communication in the workplace. SoAR calls support improving situation awareness when meeting relevant coworkers or contacts in person is not possible: when a worker must remain at her post or when it is dangerous to leave a situation unattended, when physical distance or the timeframe of the task makes co-located collaboration impossible. Whenever face-to-face collaboration serves the function of lowering the threshold to initiate communication in future situations, mediated communication is not an optimal choice, however.

The users used SoAR independently during the field tests, apart from the educational setting in BauABC, where I facilitated the test sessions. The users made more calls when the test was facilitated, and it was possible to record reactions of the users immediately after the calls. On the other hand, independent SoAR use yielded information on why using new applications and tools is so fragile: even if SoAR was considered an easy app to use, everyday obstacles easily disrupted the test use as there was no routine or compelling need to make SoAR calls. It is necessary to be aware of such hurdles in the exploitation phase of SoAR or any similar, work-related applications.

The next research phase with SoAR should be extended use in work surroundings. One form of testing SoAR further could be engaging a group of workers who have already established communication patterns among themselves. During the test phase, the users would make SoAR calls instead of audio calls, in as many situations as possible without disturbing the normal workflow. The tests would produce qualitative information on how SoAR functions in acute, actual work situations in a larger scale. This kind of testing would require facilitation, at least initially, to encourage the users to replace audio calls with SoAR. Along with qualitative research, promoting SoAR to work communities in different fields could enable collecting quantitative data from a wide range of independent users.

Future technologies will eventually make the current SoAR prototype obsolete, but in my opinion the combination of sharing the view and pointing by gesturing will remain a part of upcoming collaboration tools. According to the estimate of Goldman Sachs, AR and VR hardware adoption will accelerate considerably only in 2020's (Bellini et al. 2016, p. 8), which means there will also be room for advances in smartphone and tablet technologies in the near future. Embedded depth-sensors and 360° cameras would be most advantageous additions to mobile devices when it comes to augmented communication applications. Depth-sensing could prove effective in giving more variety and substance to gesturing: communicating the size and distance of objects would become possible. These sensors would also enable scanning objects in 3D, movement and object recognition and more accurate AR content stabilization. 360° cameras would allow independent viewing of the conversation partner's surroundings, which would further speed up identifying task-specific issues in the physical work environment.

Studying the adoption of digital applications in organizations was not the focus of this thesis, but should research on SoAR continue, adoption requirements and challenges would be important for finding out whether SoAR has the capability to actually change communication culture in the workplace. According to Hyysalo (2009, p. 52), in the adoption phase attention must be paid to the indirect effects a product has. The users should remain committed in the long run, while the attitudes towards new applications inevitably shift over time. Visual augmented communication has definitive benefits in the light of this thesis, and utilizing it on a regular basis could lead into more frequent and direct collaboration in acute situations and in planned processes alike. Effective remote collaboration would enhance the culture of mutual assistance and allow improved sharing of knowledge acquired by individual workers, as well as mitigate differences in the professional backgrounds and vocabularies of the employees.

# REFERENCES

Aloini, D., Dulmin, R., Mininno, V. & Ponticelli, S. 2012, "Supply Chain Management: A Review of Implementation Risks in the Construction Industry", *Business Process Management Journal,* vol. 18, no. 5, pp. 735-761.

Anderson, A.H., Bard, E.G., Sotillo, C., Newlands, A. & Doherty-Sneddon, G. 1997, "Limited Visual Control of the Intelligibility of Speech in Face-to-Face Dialogue", *Perception & Psychophysics,* vol. 59, no. 4, pp. 580-592.

Apache Software Foundation 2016, *Architectural Overview of Cordova Platform – Apache Cordova,* [Online]. Available: https://cordova.apache.org/docs/en/latest/guide/overview/index.html [2016, 08/29]

Azuma, R.T. 1997, "A Survey of Augmented Reality", *Presence: Teleoperators and Virtual Environments,* vol. 6, no. 4, pp. 355-385.

Baird, K.M. & Barfield, W. 1999. "Evaluating the Effectiveness of Augmented Reality Displays for a Manual Assembly Task", *Virtual Reality*, 4(4), pp. 250-259.

Baker, M., Hansen, T., Joiner, R. & Traum, D. 1999, "The Role of Grounding in Collaborative Learning Tasks", *Collaborative Learning: Cognitive and Computational Approaches*, ed. P. Dillenbourg, Elsevier, Amsterdam, pp. 31-63.

Bauer, M., Kortuem, G. & Segall, Z. 1999, "'Where Are You Pointing At?' A Study of Remote Collaboration in a Wearable Videoconference System", *Proceedings of the 3rd IEEE International Symposium on Wearable Computers*, IEEE Computer Society, Washington DC, pp. 151-159.

Bauters, M., Burchert, J., Burchert, M., Klamma, R., Koren, I., Müller, W., Ngua, K., Nicolaescu, P., Nikkilä, L., Pejoska, J. & Purma, J. 2014, *D4.2 Layers Tools for Artefact and Mobile Layer* [Homepage of the Learning Layers project], [Online]. Available: http://goo.gl/Ijs20e [2016, 08/29].

Bekker, M.M., Olson, J.S. & Olson, G.M. 1995, "Analysis of Gestures in Face-to-Face Design Teams Provides Guidance for How to Use Groupware in Design", *Proceedings of the 1st Conference on Designing Interactive Systems: Processes, Practices, Methods & Techniques,* ACM, New York, pp. 157-166.

Bellini, H., Wei, C., Sugiyama, M., Shin, M., Alam, S. & Takayama, D. 2016, *Virtual & Augmented Reality: Understanding the Race for the Next Computing Platform* [Homepage of The Goldman Sachs Group, Inc.], [Online]. Available: http://www.goldmansachs.com/our-thinking/pages/technology-driving-innovation-folder/virtual-and-augmented-reality/report.pdf [2016, 08/22].

Biehl, J.T., Avrahami, D. & Dunnigan, A. 2015, "Not Really There: Understanding Embodied Communication Affordances in Team Perception and Participation", *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing,* ACM, New York, pp. 1567-1575.

Billinghurst, M., Weghorst, S. & Furness, T. 1998, "Shared space: An Augmented Reality Approach for Computer Supported Collaborative Work", *Virtual Reality,* vol. 3, no. 1, pp. 25-36.

Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E. & Ivkovic, M. 2010, "Augmented Reality Technologies, Systems and Applications", *Multimedia Tools and Applications,* vol. 51, no. 1, pp. 341-377.

Caudell, T.P. & Mizell, D.W. 1992, "Augmented Reality: An Application of Heads-up Display Technology to Manual Manufacturing Processes", *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences,* IEEE, Washington DC, pp. 659-669.

Clark, H.H. & Krych, M.A. 2004, "Speaking While Monitoring Addressees for Understanding", *Journal of Memory and Language,* vol. 50, no. 1, pp. 62-81.

Clark, H.H. & Wilkes-Gibbs, D. 1986, "Referring as a Collaborative Process", *Cognition,* vol. 22, no. 1, pp. 1-39.

Daly-Jones, O., Monk, A. & Watts, L. 1998, "Some Advantages of Video Conferencing over High-quality Audio Conferencing: Fluency and Awareness of Attentional Focus", *International Journal of Human-Computer Studies,* vol. 49, no. 1, pp. 21-58.

Domova, V., Vartiainen, E. & Englund, M. 2014, "Designing a Remote Video Collaboration System for Industrial Settings", *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces,* ACM, New York, pp. 229-238.

Endsley, M.R. 1995, "Toward a Theory of Situation Awareness in Dynamic Systems", *Human Factors: The Journal of the Human Factors and Ergonomics Society,* vol. 37, no. 1, pp. 32-64.

Feiner, S., Macintyre, B. & Seligmann, D. 1993, "Knowledge-based Augmented Reality", *Communications of the ACM - Special Issue on Computer Augmented Environments: Back to the Real World,* vol. 36, no. 7, pp. 53-62.

Fussell, S.R., Kraut, R.E. & Siegel, J. 2000, "Coordination of Communication: Effects of Shared Visual Context on Collaborative Work", *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work,* ACM, New York, pp. 21-30.

Fussell, S.R., Setlock, L.D. & Kraut, R.E. 2003, "Effects of Head-mounted and Scene-oriented Video Systems on Remote Collaboration on Physical Tasks", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* ACM, New York, pp. 513-520.

Fussell, S.R., Setlock, L.D., Yang, J., Ou, J., Mauer, E. & Kramer, A.D. 2004, "Gestures over Video Streams to Support Remote Collaboration on Physical Tasks", *Human-Computer Interaction,* vol. 19, no. 3, pp. 273-309.

Gauglitz, S., Nuernberger, B., Turk, M. & Höllerer, T. 2014a, "In Touch with the Remote World: Remote Collaboration with Augmented Reality Drawings and

Virtual Navigation", *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology,* ACM, New York, pp. 197-205.

Gauglitz, S., Nuernberger, B., Turk, M. & Höllerer, T. 2014b, "World-stabilized Annotations and Virtual Scene Navigation for Remote Collaboration", *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology,* ACM, New York, pp. 449-459.

Gergle, D., Kraut, R.E. & Fussell, S.R. 2013, "Using Visual Information for Grounding and Awareness in Collaborative Tasks", *Human–Computer Interaction,* vol. 28, no. 1, pp. 1-39.

Gergle, D., Kraut, R.E. & Fussell, S.R. 2004, "Action As Language in a Shared Visual Space", *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work,* ACM, New York, pp. 487-496.

Helle, S., Korhonen, S., Euranto, A., Kaustinen, M. & Lehtonen, T. 2014, "Benefits Achieved by Applying Augmented Reality Technology in Marine Industry", *13th Conference on Computer Applications and Information Technology in the Maritime Industries COMPIT'14, Redworth, UK,* Technische Universität Hamburg-Harburg, Hamburg, pp. 86-97.

Huang, W. & Alem, L. 2013, "Gesturing in the Air: Supporting Full Mobility in Remote Collaboration on Physical Tasks", *Journal of Universal Computer Science,* vol. 19, no. 8, pp. 1158-1174.

Hyysalo, S. 2009, *Käyttäjä tuotekehityksessä: Tieto, tutkimus, menetelmät,* Taideteollinen korkeakoulu, Helsinki.

Härkänen, L., Helle, S., Järvenpää, L. & Lehtonen, T. 2015, *Novel Interaction Techniques for Mobile Augmented Reality applications. A Systematic Literature Review* [Homepage of University of Turku Technical Reports, No.9], [Online]. Available: http://urn.fi/URN:ISBN:978-951-29-6214-3 [2016, 05/13].

Jo, H. & Hwang, S. 2013, “Chili: Viewpoint Control and On-video Drawing for Mobile Video Calls”, *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, ACM, New York, pp. 1425-1430.

Kato, H. & Billinghurst, M. 1999, “Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System”, *Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality,* pp. 85-94.

Kim, S., Lee, G.A. & Sakata, N. 2013, "Comparing Pointing and Drawing for Remote Collaboration", *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR),* IEEE, Washington DC, pp. 1-6.

Klinker, G., Stricker, D. & Reiners, D. 2001, "Augmented Reality for Exterior Construction Applications", *Fundamentals of Wearable Computers and Augmented Reality*, eds. W. Barfield & T. Claudell, Lawrence Erlbaum Associates, London, pp. 379-427.

Kraut, R.E., Fussell, S.R. & Siegel, J. 2003, "Visual Information as a Conversational Resource in Collaborative Physical Tasks", *Human-Computer Interaction,* vol. 18, no. 1, pp. 13-49.

Kraut, R.E., Gergle, D. & Fussell, S.R. 2002, "The Use of Visual Information in Shared Visual Spaces: Informing the Development of Virtual Co-presence", *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work,* ACM, New York, pp. 31-40.

Kunkel, N., Soechtig, S., Miniman, J. & Stauch, C. 2016, *Augmented and Virtual Reality Go To Work: Seeing Business through a Different Lens* [Homepage of Deloitte University Press], [Online]. Available: http://goo.gl/XqNvC6 [2016, 08/22].

Learning Layers 2016, *Scaling up Technologies for Informal Learning in SME Clusters,* [Online]. Available: http://learning-layers.eu/ [2016, 04/26].

Lehtinen, T. 2009, *Elementtirunkovaiheen hallinnan kehittäminen*, Master of Engineering thesis, Metropolia University of Applied Sciences.

Lehtonen, T. 2016, *MARIN2: Mobile Mixed Reality Applications for Professional Use* [Homepage of Technology Research Center, University of Turku], [Online]. Available: http://trc.utu.fi/ar/research/marin2/ [2016, 05/16].

Leinonen, T. 2010, *Designing Learning Tools: Methodological Insights,* Aalto University School of Art and Design, Helsinki.

Licoppe, C. & Morel, J. 2012, "Video-in-Interaction: 'Talking Heads' and the Multimodal Organization of Mobile and Skype Video Calls", *Research on Language and Social Interaction,* vol. 45, no. 4, pp. 399-429.

Lukosch, S., Billinghurst, M., Alem, L. & Kiyokawa, K. 2015a, "Collaboration in Augmented Reality", *Computer Supported Cooperative Work (CSCW),* vol. 24, no. 6, pp. 515-525.

Lukosch, S., Lukosch, H., Datcu, D. & Cidota, M. 2015b, "Providing Information on the Spot: Using Augmented Reality for Situational Awareness in the Security Domain", *Computer Supported Cooperative Work (CSCW),* vol. 24, no. 6, pp. 613-664.

McNeill, D. 1992, *Hand and Mind: What Gestures Reveal about Thought,* University of Chicago Press, Chicago.

Minneman, S.L. & Bly, S.A. 1991, "Managing a Trois: A Study of a Multi-user Drawing Tool in Distributed Design Work", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* ACM, New York, pp. 217-224.

Molich, R. & Dumas, J.S. 2008, "Comparative usability evaluation (CUE-4)", *Behaviour & Information Technology,* vol. 27, no. 3, pp. 263-281.

O'Malley, C., Langton, S., Anderson, A., Doherty-Sneddon, G. & Bruce, V. 1996, "Comparison of face-to-face and video-mediated interaction", *Interacting with Computers,* vol. 8, no. 2, pp. 177-192.

Ou, J., Fussell, S.R., Chen, X., Setlock, L.D. & Yang, J. 2003, "Gestural Communication over Video Stream: Supporting Multimodal Interaction for Remote Collaborative Physical Tasks", *Proceedings of the 5th International Conference on Multimodal Interfaces,* ACM, New York, pp. 242-249.

Pejoska, J., Bauters, M., Purma, J. & Leinonen, T. 2016, "Social Augmented Reality: Enhancing Context-dependent Communication and Informal Learning at Work", *British Journal of Educational Technology,* vol. 47, no. 3, pp. 474-483.

Rae, I., Mutlu, B. & Takayama, L. 2014, "Bodies in Motion: Mobility, Presence, and Task Awareness in Telepresence", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* ACM, New York, pp. 2153-2162.

Rekimoto, J. 1996, "Transvision: A Hand-Held Augmented Reality System For Collaborative Design", *Proceedings of Virtual Systems and Multimedia '96,* pp. 85-90.

Rettig, M. 1994, "Prototyping for Tiny Fingers", *Communications of the ACM,* vol. 37, no. 4, pp. 21-27.

Robert, K., Zhu, D., Huang, W., Alem, L. & Gedeon, T. 2013, "MobileHelper: Remote Guiding Using Smart Mobile Devices, Hand Gestures and Augmented Reality", *SIGGRAPH Asia 2013 Symposium on Mobile Graphics and Interactive Applications,* ACM, New York, pp. 1-5.

Robinson, J., McIntyre, A. & Blau, B. 2016, *The First Three Steps in Evaluating the Role of Head-Mounted Displays for Field Service* [Homepage of Gartner], [Online]. Available: http://goo.gl/xc1DQ9 [2016, 08/22].

Rolston, M. 2013, *Today's Phones and Tablets Will Die Out Like the PC* [Homepage of MIT Technology Review], [Online]. Available: https://www.technologyreview.com/s/516486/todays-phones-and-tablets-will-die-out-like-the-pc/ [2016, 04/26].

Schall, G., Mendez, E., Kruijff, E., Veas, E., Junghanns, S., Reitinger, B. & Schmalstieg, D. 2009, "Handheld Augmented Reality for Underground Infrastructure Visualization", *Personal and Ubiquitous Computing,* vol. 13, no. 4, pp. 281-291.

Sellen, A.J. 1995, "Remote Conversations: The Effects of Mediating Talk With Technology", *Human–Computer Interaction,* vol. 10, no. 4, pp. 401-444.

Sena, P. 2016*, How The Growth of Mixed Reality Will Change Communication, Collaboration and the Future of the Workplace* [Homepage of Tech Crunch Network], [Online]. Available: https://techcrunch.com/2016/01/30/how-the-growth-of-mixed-reality-will-change-communication-collaboration-and-the-future-of-the-workplace/ [2016, 08/22].

Szalavári, Z., Schmalstieg, D., Fuhrmann, A. & Gervautz, M. 1998, "'Studierstube': An Environment for Collaboration in Augmented Reality", *Virtual Reality,* vol. 3, no. 1, pp. 37-48.

Tang, J.C. & Minneman, S.L. 1991, "Videodraw: A Video Interface for Collaborative Drawing", *ACM Transactions on Information Systems,* vol. 9, no. 2, pp. 170-184.

Uzialko, A.C. 2016*, Businesses Are Embracing AR* [Homepage of Business News Daily], [Online]. Available: http://www.businessnewsdaily.com/9245-augmented-reality-for-business.html [2016, 08/22].

Venkatesh, V., Thong, J.Y. & Xu, X. 2012, "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology", *MIS quarterly,* vol. 36, no. 1, pp. 157-178.

WebRTC Initiative 2016, *WebRTC,* [Online]. Available: https://webrtc.org/ [2016, 06/22].

Wikipedia 2016, *Videoconferencing* [Online]. Available: https://en.wikipedia.org/wiki/Videoconferencing [2016, 06/03].

Wobbrock, J.O. & Kientz, J.A. 2016, "Research Contributions in Human-computer Interaction", *Interactions,* vol. 23, no. 3, pp. 38-44.

Woodward, C., Hakkarainen, M., Korkalo, O., Kantonen, T., Aittala, M., Rainio, K. & Kähkönen, K. 2010, "Mixed Reality for Mobile Construction Site Visualization and Communication", *Proceedings of the 10th International Conference on Construction Applications of Virtual Reality*, pp. 1-10.

Zhao, D. & Rosson, M.B. 2009, "How and Why People Twitter: The Role That Micro-blogging Plays in Informal Communication at Work", *Proceedings of the ACM 2009 International Conference on Supporting Group Work,* ACM, New York, pp. 243-252.

Zhou, F., Duh, H. & Billinghurst, M. 2008, "Trends in Augmented Reality Tracking, Interaction and Display: A Review of Ten Years of ISMAR", *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, Washington DC, pp. 193-202.

## APPENDIX 1: TECHNOLOGY STACK

By Matti Jokitulppo

**AngularJS**

Angular is an open-source framework for creating single-page web applications, currently maintained by Google. It comes packaged with all the necessary boilerplate in order for developers to quickly get started making robust, modern web applications, with features such as two-way data binding. AngularJS allows writing custom attributes into HTML interpreted and parsed by its interpreter. The reason AngularJS was used with SoAR was because the Ionic mobile framework highly suggests being used together with AngularJS.

**Apache Cordova**

Cordova is an open-source framework and toolset for creating hybrid mobile applications. Essentially this means a developer can create native-feeling applications with web technologies (HTML5/CSS/JavaScript), and Cordova handles bundling the web application into a native application that can be downloaded and run on mobile devices. This enables developers to use the same codebase to develop on multiple platforms, such as Android, iOS and Windows Phone. Cordova also offers bindings to native phone functionalities, such as the camera or contacts book, features that ordinary websites cannot generally access.

**Babel & ES2015**

Babel is a transpiler for JavaScript that is usually used for transforming code from future JavaScript standards to a format that is already understood by browsers and older JavaScript interpreters. This enables developers to use new features of JavaScript, while still maintaining backwards compatibility with older browsers and interpreters. ES2015 (ECMAScript 2015) is the current latest standard. Babel is used with the SoAR server for future proofing and ease of development.

**Bower**

Bower is a package-manager meant to be used for front-end development. With Bower, you can install dependencies through Git and GitHub

**Crosswalk**

The Crosswalk project is an open-source implementation of the native Android/iOS WebView runtime. Crosswalk allows all Android devices from 4.0 and upwards to run the latest version of the Google Chromium runtime inside of them. This allows developers to use the latest features of Chrome even on older devices, in order to use new web APIs and to improve performance. Crosswalk was chosen for SoAR for aforementioned performance improvements, and to enable WebRTC for older devices.

**Express**

Express is a popular web application framework for Node.js. It is used in the SoAR server to set up a RESTful API to manage user actions such as logging in and registering through HTTP end-points.

**Git & GitHub**

Git is a distributed version control system originally developed by Linus Torvalds to aid with development of the Linux kernel. GitHub is a popular hosting service for git repositories, but it also comes with its own added features, such as a wiki system and an issue tracker.

**Gulp**

Gulp is a streaming build-system meant for web applications. It helps automate menial tasks such as compiling SASS to CSS, removing extra characters from JavaScript and HTML files for a smaller final file size and automatic refreshing of the browser on file changes.

**Ionic**

Ionic is a framework for assisting in creating hybrid mobile applications together with Cordova. It comes packaged with its own components and helpers for things such as mobile gestures and touch events. Ionic is coupled together with AngularJS, but one can use the stylized components without it.

**JavaScript**

JavaScript is an interpreted scripting language used in creating interactive web experiences. It is currently based on the ECMAScript standard. It was originally meant to be run inside ordinary web browsers, but nowadays it can be used to make mobile apps, desktop software and server side tools.

**Node.js & NPM**

Node.js is a runtime environment for creating server-side web applications using JavaScript. Node.js was chosen for SoAR due to its speed and similarity with client-side programming. NPM (Node package manager) is the package manager for Node.js modules.

**Postgres**

Postgres (or PostgreSQL) is a robust open-source enterprise-grade relational database.

**SASS**

SASS (Syntastically Awesome Stylesheets) is a scripting language that extends and adds feature to CSS (Cascading Style Sheet) with features such as variables, nested CSS selectors and functions. Before being interpreted by the browser, SASS files need to be first compiled down to CSS.

**WebRTC & PeerJS**

WebRTC (Web Real-Time Communication) is a browser API to enable serverless browser-to-browser communication, in the form of voice and video chat or P2P file sharing, without the usage of browser plugins such as Flash. WebRTC plays a very crucial part in SoAR in handling the actual user-to-user video call functionality

PeerJS is a JavaScript library that wraps the native WebRTC features into a more usable API for ease of usage.

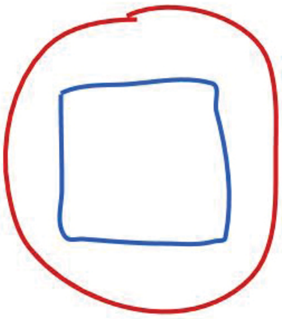
**WebSockets & socket.io**

WebSockets is a protocol for enabling a two-way communication channel client and server. Before WebSockets, real time web applications had to be implemented with periodically polling the server for new data, which is quite inefficient. Socket.io is a wrapper library around native WebSockets that further standardizes and enhances their functionality.

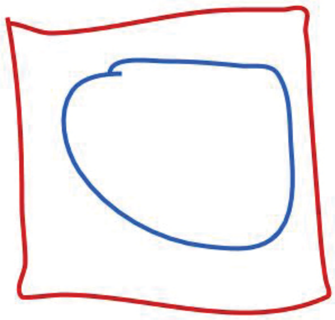
# APPENDIX 2: USABILITY TEST TASKS

**First, let's test drawing on the screen together. Switch to fullscreen if needed.**

1. Draw this together on YOUR camera view:



2. Draw this together on your PAIR'S camera view:

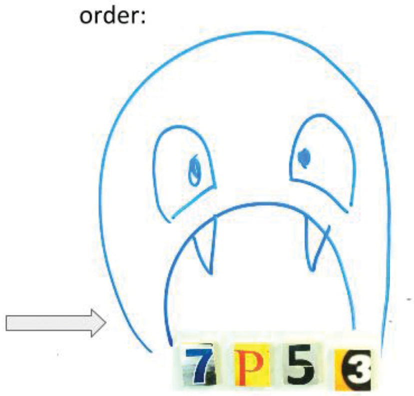


Find a whiteboard and a blue marker in the room.

Draw this face on the board with the marker:

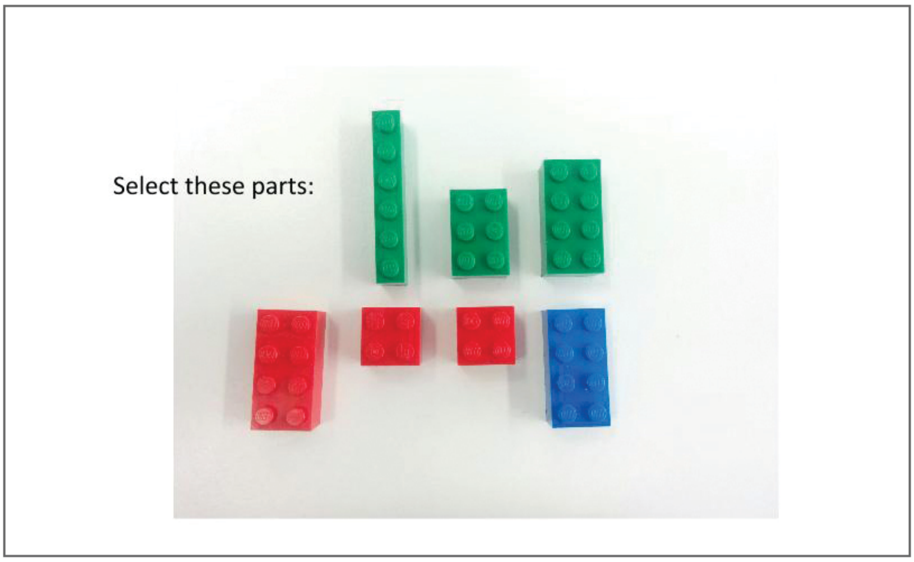
Find these characters in the room and pick the circled ones.

Place them on the drawing in this order:

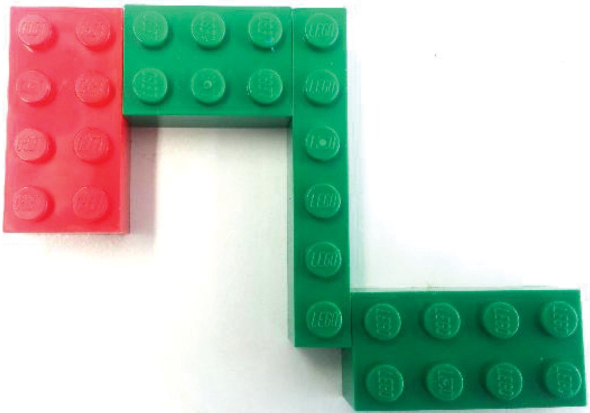


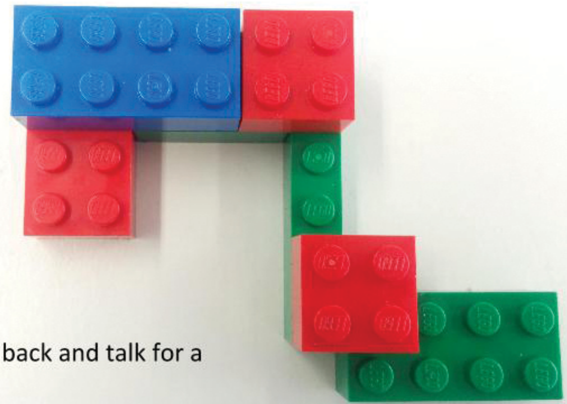

The stuff for the
next task is here:
X


This should go on a table somewhere.


Select these parts:

Start building:





Done?
That’s it, come back and talk for a while.