



# shARe-IT: Ad hoc Remote Troubleshooting through Augmented Reality

Thomas Ludwig<sup>1\*</sup> , Oliver Stickel<sup>2</sup>, Peter Tolmie<sup>3</sup> & Malte Sellmer<sup>1</sup>

<sup>1</sup>*Cyber-Physical Systems, University of Siegen, Kohlbedtstr. 15, 57072 Siegen, Germany (E-mail: thomas.ludwig@uni-siegen.de);* <sup>2</sup>*Ruhr-University Bochum, Bochum, Germany;* <sup>3</sup>*University of Siegen, Information Systems and New Media, Siegen, Germany*

Accepted: 21 January 2021

**Abstract.** 10 years ago, Castellani et al. (Journal of Computer Supported Cooperative Work, vol. 18, no. 2–3, 2009, pp. 199–227, 2009) showed that using just an audio channel for remote troubleshooting can lead to a range of problems and already envisioned a future in which augmented reality (AR) could solve many of these issues. In the meantime, AR technologies have found their way into our everyday lives and using such technologies to support remote collaboration has been widely studied within the fields of Human-Computer Interaction and Computer-Supported Cooperative Work. In this paper, we contribute to this body of research by reporting on an extensive empirical study within a Fab Lab of troubleshooting and expertise sharing and the potential relevance of articulation work to their realization. Based on the findings of this study, we derived design challenges that led to an AR-based concept, implemented as a HoloLens application, called shARe-it. This application is designed to support remote troubleshooting and expertise sharing through different communication channels and AR-based interaction modalities. Early testing of the application revealed that novel interaction modalities such as AR-based markers and drawings play only a minor role in remote collaboration due to various limiting factors. Instead, the transmission of a shared view and especially arriving at a shared understanding of the situation as a prerequisite for articulation work continue to be the decisive factors in remote troubleshooting.

**Keywords:** Augmented reality, Remote troubleshooting, Expertise sharing, Fab lab, CSCW

## 1. Introduction

Problems arise in many situations in life. Sometimes we cannot solely resolve these problems on our own, prompting us to ask for assistance from people with a more comprehensive or technical expertise (Crabtree et al. 2006; O’Neill et al. 2005). Whether the problems pertain to bug tracking within software support centres (McDonald and Ackerman 1998), customers’ troubleshooting issues with their office devices (Castellani et al. 2009), medical diagnosis within hospitals (Cicourel 1990), or situation assessment within crisis control rooms (Ley et al. 2014), we often encounter practical troubles that simply exceed our own capacity.

In the past, seeking technical assistance entailed an expert's physical presence to undertake the troubleshooting (Crabtree et al. 2006). For example, electrical service engineers had to go to offices and worksites to repair broken printers (Orr 1996). This type of practice is expensive (O'Neill et al. 2005). Bowers and Martin (2000) highlight the concept of "new factories," which roughly describes telephone call centres for outsourcing engineering support. This concept refers to the differentiation between classic industrial factories that focus on manufacturing goods for the customer and new factories that provide a more sophisticated service for the customer "through assembling an army of call-takers each at a terminal to the corporate database" (Bowers and Martin 2000). Such a service might involve offering some form of support without the need to visit the site, thus saving time and costs and enabling experts to assist more help-seekers.

Although these new factories and remote troubleshooting have several advantages in terms of time and cost efficiency for the help-giver and the help-seeker, they also have considerable downsides. Referring to the case of network printing technology, Castellani et al. (2009) revealed numerous dislocations during remote troubleshooting processes, which complicate cooperation and expertise sharing between a remote troubleshooter and a help-seeker. The troubleshooter has to rely on the help-seeker to clearly articulate the problem, while the expert in turn has to verbally describe the actions that the help-seeker has to perform in a language she understands. This approach requires a significant amount of articulation work and a degree of "recipient design", in a context characterized by inadequate mutual knowledge (Crabtree et al. 2006; Whalen and Vinkhuyzen 2000). Put simply, remote troubleshooting is not merely technical (Orr 1996). When trying to support remote troubleshooting with information and communications technology (ICT), functionality should therefore utilize "computer-based mechanisms that support articulation work" (Crabtree et al. 2006).

More than 10 years ago, Castellani et al. (2009) concluded that, in the case of remote troubleshooting, "many of the problems related to the audio channel." In addition to their approach of turning the device itself into a knowledge base, the authors described various approaches to tackling the issue with augmented reality (AR). Augmented reality enhances the features of the real world with multi-modal virtual information. The concept has found its way into readily available consumer products such as Google Glass and the Microsoft HoloLens. These technologies have the potential to enhance verbal communication through a visual channel, which could support conversational grounding (Bauer et al. 1999; Clark and Marshall 1981) and therefore collaboration (Lukosch et al. 2015).

We acknowledge that using concepts from AR within collaborative settings is not new. Within this paper, however, we intend to build upon the existing discourse in human-computer interaction (HCI) and especially its subarea of computer-supported cooperative work (CSCW) regarding remote troubleshooting and expertise sharing by exploring new AR technology such as smart glasses, to tackle several types of dislocation (Castellani et al. 2009). The paper contributes both a design case study

(Wulf et al. 2015), that was undertaken to explore the opportunities, and a discussion of the limitations of current AR technologies for supporting remote troubleshooting. In the next section, we analyse related approaches to (remote) troubleshooting as well as expertise sharing and how to support troubleshooting with IT. We then report on a qualitative empirical study that sought to explore the impact of expertise sharing and troubleshooting within a complex, evolving, ICT-driven application domain (3D printing). Our goal was to understand how troubleshooting works in practice and what constitutes “good” troubleshooting or expertise sharing, without any dependency on any particular modality. We therefore provide empirically-grounded descriptions of expertise-sharing practices in a non-traditional and interdisciplinary organization (a fabrication laboratory or Fab Lab). Based on these practices, we derived an approach that allows for remote troubleshooting and guidance based on AR. We used the empirical findings and existing findings in the literature to develop a HoloLens-based prototype, shARe-it, which was designed to support remote troubleshooting between experts and non-experts via AR. A basic implementation of the prototype within a laboratory revealed that, although AR can support current troubleshooting processes, it can only be viewed as a supplement to the verbal negotiation present in interpersonal processes and that the immaturity of current AR technology results in a set of ongoing challenges, including: a need for improved interleaving between AR technology and other technical resources, such as sensors; a need for improved accuracy in existing resources, such as markers; a continuing need for improved deictic resources; a need for better ways of capturing a relevant and variable visual field, without environmental interference; a continuing reliance upon verbal description to make sense of what is being seen and what should actually be done; and a continuing need to arrive, in situ, at a mutually comprehensible terminology.

Our contribution to the CSCW community is threefold (Wobbrock and Kientz 2016): 1) An empirical contribution that gives us insights into troubleshooting practices in a complex hardware-oriented application domain, with a focus on 3D printing; 2) A report of the design and basic testing of an AR-based application for smart glasses to support ad hoc and remote troubleshooting practices in order to reduce the kinds of dislocation first described by Castellani et al. (2009); 3) Identification of the current hardware shortcomings of smart glasses with regard to supporting these kinds of practices.

## 2. Related work

CSCW has always had a strong focus on the social practices involved in expertise and knowledge sharing, as well as designing innovative ICT artefacts that could support these types of activities (Ackerman et al. 2013). The key concern is that ICT artefacts must combine the “information space (knowledge artefacts) with the communication space (social interactions)” (Pipek et al. 2012), which is a key innovation for knowledge management (Ackerman and Malone 1990; Huysman and Wulf 2005).

Distinctions have emerged between knowledge and expertise sharing in the CSCW discourse. The term “knowledge sharing” adopts a perspective in which the “externalization of knowledge in the form of computational or information technology artefacts or repositories play an important role. We use the term “expertise sharing” when the capability to get the work done or to solve a problem is instead based on discussions among knowledgeable actors and less significantly supported by a priori externalizations” (Ackerman et al. 2013). Dengel (2016) explains the distinction between knowledge and expertise sharing in this manner: “While knowledge means to memorize the right things, expertise means to do the things right and even anticipate what might be the right thing to do.” Kristoffersen et al. (1996) also situate expertise in the associated social activities that are performed to accomplish a related task. These perspectives can also be seen to relate to older discussions regarding the distinction between “knowing how” and “knowing that” (Ryle 1945).

### 2.1. Remote expertise sharing and troubleshooting

There are different types of troubleshooting. In one type, the unavailability of a (remote) expert prompts the help-seeker to proceed with the activity on her own. The challenge in this case is providing the appropriate resources needed by the help-seeker to detect and solve the problem at hand. Current approaches seek to enhance machines by incorporating troubleshooting functionality (Castellani et al. 2009; Crabtree et al. 2006; Hoffmann et al. 2019; Ludwig et al. 2017, 2019). Another type of troubleshooting encompasses situations in which a remote expert is available. The challenges in this case involve providing the help-giver with the appropriate information about the ailing artefact to enable her to identify the problem and then communicating the actions needed to solve the problem to the help-seeker (Castellani et al. 2009). In this paper, we focus on the second case. Castellani et al. (2009) describe the process of remote troubleshooting and its shortcomings as follows:

1. The remote expert receives a problem description from the help-seeker, which tends to be incomplete.
2. The expert and help-seeker seek to narrow down the problem through collaborative interaction to enable the expert to derive possible solutions. During this process, the help-seeker might need to conduct some tests to provide the expert with additional information.
3. The actual troubleshooting process occurs when the expert gives instructions and guidance to the help-seeker, who has to undertake them and report the results back to the expert.

The first two steps indicate that troubleshooting increases in difficulty if a help-seeker is unable to articulate the problem (Ludwig et al. 2014). The expert has to transform the help-seeker’s description of the problem into “a more specific and technical language” (Crabtree et al. 2006), before being able to search and identify

possible solutions. The expert then has to translate a technical and specific solution to a problem into a language that the help-seeker understands. These different factors are grouped into the following three categories of dislocation (Castellani et al. 2009):

1. Physical dislocation solely occurs in remote troubleshooting situations. The physical dislocation focuses on the physical distance between the place in which the problem occurs and the location of the expert who is able to provide support. This physical dislocation results in the expert not being able to diagnose and fix the problem by herself and instead needing to collaborate with the help-seeker. The expert has to rely on the verbal feedback of the help-seeker and the description of the symptoms and reactions of the ailing artefact. The expert then has to describe the actions that the help-seeker should perform.

2. Conceptual dislocation is based on the difference in the levels of knowledge between the help-seeker and the expert. This form of dislocation creates problem descriptions that may on the one hand include unnecessary detail and on the other hand be lacking relevant information. As Castellani et al. (2009) underscore, “the relevancies for the customer and the troubleshooter are distinct and they need to arrive at a mutual understanding.” Establishing a “mutual understanding” is a major part of troubleshooting. Repair work is a feature of misunderstandings or lack of understanding in conversational exchanges in general and is arguably often generated by problems of “formulation” (Whalen et al. 1988; Whalen and Vinkhuyzen 2000).

3. Logical dislocation focuses on being able to consider the history and context of an ailing device that needs troubleshooting. This might not be fully applicable in all troubleshooting situations. Often help-seekers do not have all the historical and contextual information available for the expert, thus hindering the troubleshooting process.

## 2.2. ICT-supported remote troubleshooting and expertise sharing

Early systems for remote collaboration used 2D monitors or handheld devices in combination with video cameras to share the view between two different workspaces (Gao et al. 2016). These remote collaboration systems focused on adding visual aspects such as annotations and hand gestures.

In a system geared towards remote guidance, Gao et al. (2016) provide an oriented point-cloud view in a mixed reality setting to enhance the “spatial awareness for the users of remote collaboration system[s].” The expert and worker must wear a virtual reality (VR) headset with an attached depth sensor. On the worker’s side, the view is captured and provided to the expert. On the expert’s side, both hands are captured and provided as a point-cloud in both views. The worker has the guiding hand gestures of the expert superimposed on her view of the workplace and, by using the VR headset, the worker is able to receive the expert’s instructions. Tecchia et al. (2012) present an approach that only needs the remote expert to wear a VR headset. A 3D camera is positioned above the worker’s workspace that is equipped with a

monitor offering a view of the physical objects, her own hands, and the superimposed hands of the remote expert (Figure 1). Although this setup aims to provide a “shared virtual interaction space” (Tecchia et al. 2012), it is limited in its scope to support help-seekers in mobile settings.

Adcock et al. (2013, 2014) use depth cameras to perform real-time 3D position capture of workers and their environment. These data are streamed to an expert on a multi-touch monitor where they can adjust their own view on the workspace. To guide the remote worker, the expert can use finger gestures to draw on the scene. These drawings are directly projected onto the workspace, utilizing a projector placed above the scene. Beyond desktop workspaces, several approaches focus on remote troubleshooting in mobile settings. Huang and Alem (2011) concentrate on mobile workers and provide a near-eye display in combination with a camera that is attached to a worker’s helmet. The expert uses a touchscreen-enabled monitor that displays the video from the worker’s camera as a shared virtual space. Usability evaluation of this system generated very positive results, suggesting that systems can be improved by making cameras match a worker’s view.

Nowadays, augmented reality is often used for remote troubleshooting. Augmented reality is a concept that is designed to enrich reality with the help of artificial and virtual elements by overlaying virtual images within the real world. Milgram and Kishino (1994) present the mixed reality continuum that describes the different variations of combining virtual and real elements, with a real environment sitting at one end of the spectrum and a virtual environment sitting at the other, and various mixtures of augmented reality and augmented virtuality in-between. Beyond the differentiation between reality and virtuality, Azuma (1997) provides a definition for AR systems that includes two additional characteristics, namely “interactive in real time” and “registered in 3D.” “Interactive in real time” sets AR apart from movies or similar forms of presentation of content in which real elements are combined with virtual ones in advance, thus preventing there being any type of interactive influence over the content.

Early approaches focused on fostering mobile face-to-face collaboration through AR (Henrysson et al. 2005). Gauglitz et al. (2014) developed a prototype that is based on a tablet device and supports a remote user in exploring the on-site scene independently of the local user’s current camera

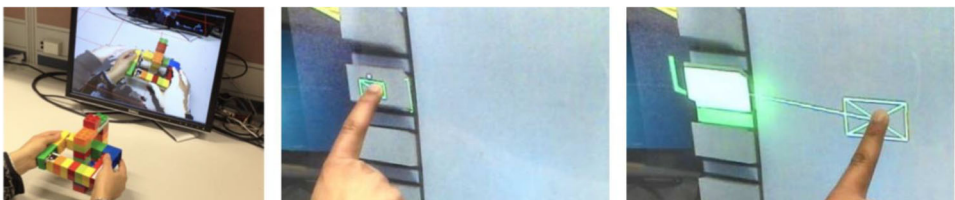


Figure 1. Shared virtual interaction space (Tecchia et al. 2012) (left); selecting an object (Adcock et al. 2014) (middle); moving an object (Adcock et al. 2014) (right)



position. The remote user is also able to communicate with the local user via spatial live annotations that are visible through augmented reality. This study shows how remote collaboration can significantly benefit from bridging between classical video conferencing systems and remote world exploration. Gavish et al. (2011) utilize a tablet PC equipped with a camera as a display that is mounted on a movable stand to allow adjustment of the view while still keeping the user's hands free. The system provides training for assembly tasks and enables the trainee to work on a real device while seeing the instructions directly superimposed on the concerned parts. Adcock et al. (2014) have also presented a spatial augmented reality (SAR) approach that projects instructions from a remote expert with light on the workspace, thus enabling multiple users to simultaneously observe them.

In later work, Adcock and Gunn (2015) developed a spatial hands-free AR-based prototype to guide an on-site worker via a pico-projector on top of a helmet coupled with a camera. The camera image is streamed to another laptop where an expert can draw annotations that complement the voice communication. Other approaches have focused on camera-equipped tele-operated robotic arms to monitor the hands of the help-seeker (Gurevich et al. 2012). Poelman et al. (2012) use head-mounted displays for supporting remote collaboration between experts in crime scene investigations. Oda et al. (2013), meanwhile, focus on see-through, head-worn displays for experts and local users to navigate through pre-recorded viewpoints of a local site. Wang et al. (2019) developed a novel mixed reality-based collaborative platform for manufacturing tasks, which projects remote expert's gestures into the actual worksite to improve performance and generate a sense of co-presence. Evaluation of this platform confirmed improvements in performance during the tasks. However, the system does not offer specific support for mobile settings.

When focusing on head-mounted displays for supporting remote collaboration, Tait and Billinghurst (2015) have shown that full-view independence for remote participants increases the speed and quality of collaborative tasks, but does not improve task accuracy. Unfortunately, in the context of remote troubleshooting, improved accuracy is urgently needed. Tait and Billinghurst (2015) also noted that partial-view independence, for instance, remote users being able to freeze their view, does not benefit remote collaboration. On the other hand, Fakourfar et al. (2016) suggest that automatically freezing the video while drawing annotations can be surprisingly effective in facilitating the completion of remote assistance tasks.

As the literature shows, there is a variety of approaches to implementing AR for remote collaboration. However, the notion of using AR alone does not specify the kind of hardware that should be used. Even utilizing visual information is not mandatory, with audio or haptic information being equally possible. Nonetheless, mapping visual information into the user's view is widely used and makes sense for the majority of approaches.

### 2.3. Techniques for remote collaboration

As we have seen, Adcock et al. (2014) and Wang et al. (2019) use projectors to display the manual instructions given by an expert. Gao et al. (2016), meanwhile, present the hands of the expert as a point-cloud to the worker in their shared virtual space. Tecchia et al. (2012) fuse the captured hands of the expert together with the video feed from the workspace and Huang and Alem (2011) capture the hands of the expert from the workspace and send this back to the worker. Gauglitz et al. (2014) use a tablet device to overlay virtual content from a remote expert on the physical world and Tait and Billingham (2015), as well as Poelman et al. (2012), use head-mounted displays to provide a mobile, hands-free approach. In all of these situations, the expert and the help-seeker need to find a common ground for communication, something that is fostered in situ by linguistic and physical co-presence (Clark and Marshall 1981).

Fussell et al. (2004) noted that people use different types of gestures to underpin their spoken words. They therefore introduced a pattern of conversational grounding during collaborative tasks, which focused on three types of tasks, accompanied by a matching form of gesture. The first type of task was object identification, in which “collaborators come to mutual agreement upon or ‘ground’ the objects to be manipulated using one or more referential expressions” (Fussell et al. 2004). Pointing or deictic gestures are mostly likely to be used for this kind of task. The second type involved providing procedural statements, which offered “instructions for procedures to be performed on those objects” (Fussell et al. 2004). Instructions are mostly likely to be accompanied by representational gestures, as these gestures are used to represent the desired outcome of an action or the action itself to be performed. The third type of task is monitoring comprehension and/or task status. In this task, the expert and help-seeker “check [the] task status to ensure that the actions have had the desired effect” (Fussell et al. 2004). Henderson and Feiner (2011) conclude that directly superimposing instructions on the objects themselves (Poelman et al. 2012; Tait and Billingham 2015), or near them, is more efficient than presenting them on a nearby display (Adcock et al. 2014; Gao et al. 2016; Huang and Alem 2011; Tecchia et al. 2012).

By comparing the surrogates of pointing and drawing gestures, Fussell et al. (2004) highlight that the collaborators’ ability to maximize the use of both pointing and representational gestures is important for grounding their conversations. Fussell et al.’s study indicates an efficient use of verbal communication in connection with drawing-based interaction because of the wide use of deictic vocabulary such as “this” and “that,” which the authors claim was “undistinguishable from the use of deixis in the side-by-side conditions” (Fussell et al. 2004) they had observed in previous studies (Fussell et al. 2000). This point strengthened their argument that the “asymmetry between workers’ and helpers’ ability to point within the shared visual field” is a critical aspect of performing a task. Thus, eliminating this asymmetry by



enabling both sides to use gestures in shared spaces results in improvements in performance.

Kirk and Fraser (2006) argue for a mixed ecologies approach “advocating the use of unmediated representations of hands as the gesturing format and the presentation of gestures projected into worker’s task spaces.” To evaluate their mixed ecologies approach, the authors compared hands-only gestures with hands and sketching gestures and digital sketches. The results suggest that the use of hands-only gestures is the most efficient means of conveying actions. A combination of hands and sketching gestures resulted in a poorer performance of their set task, and pure sketches proved to be the least effective of all. One of the major constraints that hampers efficiency is that the worker has a limited awareness of what the expert sees (Kirk and Fraser 2006). Despite this finding, Kim et al. (2019) suggest that a combination of sketches and hand gestures is the most efficient mode of conveying actions when manual tasks are guided remotely. Thus, there is not a clear consensus about the most appropriate deictic mechanisms to use to support remote collaboration.

### 3. Research gap and approach

Supporting troubleshooting and expertise sharing within remote situations has been heavily studied within the CSCW community. As the literature shows, articulation work is an essential part of remote troubleshooting. However, how best to accomplish it remains unclear. Castellani et al. (2009), specifically emphasize that troubleshooting “still suffers from many of the problems related to the audio channel”. Thus, in addition to turning the device itself into a knowledge base (cp. (Ludwig et al. 2017)), the authors suggest there is scope to use AR approaches to tackle these issues.

Today, AR has been used to develop many different approaches for remote troubleshooting and expertise sharing. Most of the existing approaches entirely focus on gestures and conveying intended interaction through visual guidance. However, the sharing of resources and information, which is one of the key aspects involved in reaching a mutual understanding between the expert and the help-seeker (Crabtree et al. 2006), is not only limited to gestures. Giving background information in addition to the “how-to-do-it” knowledge (Gavish et al. 2011) supports not only remote troubleshooting but also expertise sharing and may enable a help-seeker to act more independently in the future.

Early studies in the field of AR-supported remote troubleshooting encompass the installation of a camera combined with monitors, projectors, or VR headsets. Using this setup means that establishing the environment for remote guidance is cumbersome and places various constraints upon the physical position and environment within which it can be realized (e.g., the positioning of the projector or the viewing angle of the camera). Shortly afterwards, mobile-based approaches became available (Tecchia et al. 2012). However, these did not superimpose information on the real

environment and often still relied upon cumbersome equipment (Henderson and Feiner 2011), hampering the scope for ad-hoc troubleshooting. Since then, AR technologies have developed rapidly. Nowadays, modern smart glasses like Google Glass or the Microsoft HoloLens allow off-the-shelf solutions for a hands-free use of augmented information projected directly into a user's field of view. Based on these developments, there are already approaches that encompass AR to support remote troubleshooting settings (Ens et al. 2019; Kim et al. 2018, 2019).

In this paper, we want to expand upon the current state-of-the-art regarding AR-based remote troubleshooting by adding a practical perspective (Kuutti and Bannon 2014). Here, we focus on how current off-the-shelf technologies and their characteristics affect the dislocations originally uncovered by Castellani et al. (2009). In so doing, we contribute insights into the potential and challenges of modern AR technologies for remote troubleshooting and expertise sharing. In particular, we examine what the artefacts could look like and what type of interactions these artefacts might allow (Crabtree et al. 2006). Thus, we not only contribute a comprehensive empirical study of expertise sharing and troubleshooting, but also empirically-based design implications to inform the future development of AR applications for remote expertise sharing and remote troubleshooting.

As Normark and Randall (2005) have previously stated, a system grounded on analysis of current work practice tends to be considerably more useful because it “[does] not ignore the social aspects of knowledge sharing” (Ackerman et al. 2013). One of our challenges was therefore to first understand the field of practice in order to understand the set of potentially relevant troubleshooting and expertise sharing practices. Our study was therefore organized into three different steps, following the well-established practice-based design case studies approach (Wulf et al. 2015). This approach can be broadly broken down into: an empirical pre-study; a design phase; and a subsequent evaluation of the implemented application. This approach is broadly aligned with the European CSCW tradition of grounding systems design in practice.

We first conducted an empirical-based pre-study to examine the complexity of troubleshooting and expertise sharing. We selected 3D-printing as our research domain because it exemplifies a modern, highly digitized approach to production that is not yet well understood in its application and development potential. Hence, many concerns, different use cases, and operative practices emerge, as well as ensuing challenges for users, resulting in a significant requirement for expertise sharing. We therefore chose a local Fabrication Laboratory (Fab Lab). Of course, it is important to note that our pre-study is a co-located setting in which troubleshooting and expertise sharing takes place locally. However, our goal in the pre-study was to examine a setting in which experts and inexperienced users work closely together to understand how expertise is shared and guidance provided under relatively ideal conditions. This involved holding in abeyance studies of authentic remote settings (e.g. call centres or support hotlines) to prioritize in situ expertise sharing. This, we felt, would give us insights regarding what kinds of practice remote support would

need to properly attend to. We therefore always had an eye on how the co-located exchange of experiences and the local articulation of problems and possible solutions could be transferred to possible remote troubleshooting settings. Some older work (Orr 1996; Whalen et al. 1988; Yamauchi et al. 2003) focused on in situ guidance of this order and this was, in turn, used in some subsequent studies, e.g. O'Neill et al. (2005) to inform aspects of the solutions explored. However, modern complex and technology rich domains, such as that exemplified by 3D printing, are at some remove from these original settings, and some of the assumptions made about technology and the practices adopted have inevitably evolved. This goes as much for help-seekers as it does for experts.

The design case study within this paper is a follow-up to a study published in Ludwig et al. (2014), in which we mainly reported on insights into people's reasons for engaging in 3D-printing and the types of challenges that can arise during this process. In the previous paper (Ludwig et al. 2014), we uncovered a lot of different problems and practices involved in coping with the challenge of end-users appropriating 3D printers when examining two semi-professional communities. The main challenges relevant to the current paper and its interest in troubleshooting were that: (a) identifying and locating problems is a problem itself because of the highly context-dependent nature of problems when they arise; (b) inexperienced users have difficulty in asking for help because there is often not enough contextual information available; (c) knowledge, tips and hints in sharing communities are scattered here and there, without any unifying structure, so they are not really searchable; and (d) community-specific terminology hampers the appropriation of 3D-printing.

This paper builds upon the state-of-the-art in three specific ways. First, it reports on a set of theoretical and empirical insights acquired by conducting an empirical investigation of expertise sharing by and between humans within a complex, evolving application domain. Second, it reports on the conceptual foundation and development of a prototype AR-based support system for such applications and domains and testing of its basic functionality. Third, it outlines the challenges associated with using new kinds of AR technologies when it comes to supporting remote troubleshooting and expertise sharing.

#### **4. Empirical pre-study**

We conducted our empirical pre-study within a Fab Lab (Figure 2). Fab Labs are a global network of similarly equipped and connected tech workshops that are accessible to anyone and utilized by a variety of users for everything from rapid prototyping for business to teaching, research and hobbies. The open, interdisciplinary, hierarchically flat, and not directly market-oriented structure of such a lab is certainly not representative of average organizations. Nonetheless, local Fab Labs can still provide a uniquely diverse mix of demography, application domains and user motivations, albeit with less organizational constraints. This allows for an examination of how different kinds of users actually make new, digitized production

machines work for their specific purposes. The goal of our pre-study was to understand what exactly constitutes troubleshooting in such complex settings, the kinds of practices used, and if and how expertise sharing takes place. Various ethnographic and participatory studies have been previously conducted in the local Fab Lab. Constant “background” research activities are also undertaken, such as the documentation of ongoing projects, recording of regular field notes, and extensive documentation of photos. Although the findings and data from those activities informed the research we report on in this contribution, the main data with which we concern ourselves here emanates from a new pre-study, as reported below (see 4.1).

#### 4.1. Methodology

First, we conducted an observational study within the Fab Lab by taking handwritten field notes, followed by an interview study with two experts. The experts were the coordinators of the Fab Lab and were always on site. Therefore, they had the best overview of current conditions within the Fab Lab and how expertise sharing was structured. Since there were only these two Fab Lab coordinators in total, they were the only ones interviewed. The interviews lasted about 90 min each and were audio-recorded and transcribed using a relatively pragmatic approach, where only immediately obvious anomalies in speech or nonverbal patterns were annotated. Both the transcriptions and field notes from the observational phase were originally documented in German. Data from all phases were then analysed using a thematic content analysis approach.

For the first phase, we observed people within the Fab Lab undertaking various activities during so-called Open Lab sessions. These sessions take place every Friday when the lab is open to anybody who is interested (not only students, but also companies or private individuals). During this time, the people present have free access to the Fab Lab’s resources. The wide range of machines and users make expertise sharing, troubleshooting, and other support activities in the Fab Lab highly contextualized (see Ludwig et al. (2017) and Ludwig et al. (2014) for the contextual factors of 3D printing). Hence, the entire Fab Lab community and especially the Open Lab days are based around and constantly build upon knowledge and expertise

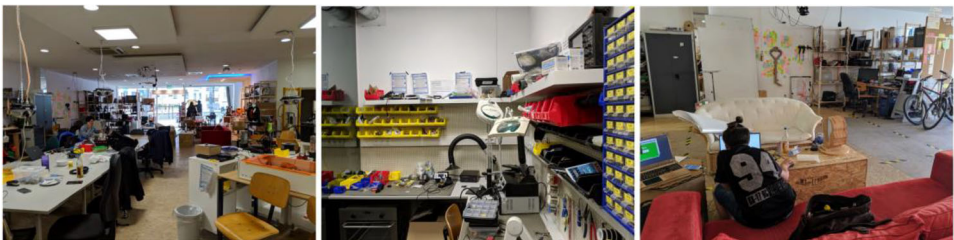


Figure 2. Fab Lab Siegen

sharing between its members. This fits our interest in analysing the practices of troubleshooting and expertise sharing.

In total, we attended six full Open Lab days, each lasting four hours (16 h in total). We engaged with the community members in a friendly and open fashion, which was consistent with the overall perceived culture during Open Labs. All the community members who have undergone safety instruction (which is mandatory before using the machines) have been informed about and consented to the Fab Lab's general position as a research infrastructure in the University and the attendance of researchers who are interested in the lab, its users, and their projects on a general level. In situations where certain individuals, groups, or projects are engaged with using in-depth methods such as in-situ interviews, or where photos are taken of the concerned individuals, they are individually informed about the study and their data is only used if they have given their explicit consent. Such in-depth engagements mostly transpire around the aforementioned safety instructions, or in situations where problems, questions, or other opportunities for troubleshooting and expertise sharing arise during Open Lab sessions. The latter instances typically became noticeable to us when individuals or groups asked around for advice, or were given increased attention by the supervising staff members (Fab Lab coordinators). We also had an interest in ad-hoc workshops or experiments and instances of "over-the-shoulder learning."

We took field notes and photos during the observations themselves, focusing on the existing tools, media, and their usage (Wulf et al. 2015). We made a conscious decision not to record videos as the Fab Lab, with a total area of 300 m<sup>2</sup> (approx. 359 yd<sup>2</sup>) and a variety of different machines, offers too many possibilities for interaction that would not have been possible to capture effectively. Instead, we participated, observed and took field notes every time communication took place between visitors, explanations were offered, or there were any kinds of exchanges of experience.

During our observations, we still tried to note as many gestures as possible because they had already been identified as an important part of the grounding process during (co-located) conversations and collaboration (Fussell et al. 2004). We also focused upon the types of information shared: descriptions of physical actions, functionality of machines, processes at a higher level, accounts of previous experience relating to individual projects, or concrete directions and step-by-step instructions. The interactional setting and the media that were used were also recorded. We did not, however, concern ourselves with personal or specific project details beyond what was necessary to understand specific situations.

The purpose of the later, semi-structured interviews with experts was to improve our understanding of the practices previously detected during our observations in the Fab Lab. Within this study, we focused only on interviews with experts, as the results of interviews with people seeking help had already been published in a previous paper (Ludwig et al. 2014). Within the expert interviews we intended to obtain a better understanding of the social context of the interactions that unfold in this community. We interviewed two of the "Fab Lab coordinators" (I01, I02), who are

University employees that are responsible for managing and running the Fab Lab as an inter-disciplinary organization. Since our study, the group of coordinators has been expanded. The interview questions were informed by the observations we had made during our previous visits and focused on the following points:

1. A general background encompassing prior experience with related technologies and communities;
2. How the interviewees got involved in the Fab Lab and digital fabrication and their personal motivations;
3. The interviewees' personal appropriation process regarding 3D printing, how they obtained their current knowledge and expertise, and the people, media, or tools that they found especially helpful;
4. The interviewees' method of handling co-located support in the Fab Lab and the particular tools and media they use;
5. How giving help is handled inside the Fab Lab community, again from the perspective of the interviewees, with a focus on tools, media, or practices;
6. The (perceived) most important elements for help-giving that should be present for remote support systems.

## 4.2. Results

In the following sections, we report on the aggregated results of the observational and interview studies. Illustrative excerpts from the transcripts and observational notes supplement the respective results and findings.

### 4.2.1. *Terminology and conceptual barriers*

As is likely the case with most communities, community-specific vocabulary and certain shared concepts exist amongst Fab Lab participants, 3D printer users, and other (sub-)groups. Both interviewees indicated that being physically co-located within the Fab Lab and the resulting opportunities to visualize and directly demonstrate things, were beneficial for overcoming unfamiliar terminology and concepts with users. In the case of (primarily) online communities, these barriers have been found to be potentially more difficult to overcome (cp. (Castellani et al. 2009)). However, digital sharing and data, heavy use of non-professional tutorials, and discussions on the internet and other remote activities were emphasized as crucial and even partly responsible for the existence of a global Fab Lab network (cp. (Ludwig et al. 2014)). As the Fab Lab coordinators have many things to do at the same time during Open Lab sessions and cannot always support individuals with their problems, I02 recommended a visual troubleshooting guide to trace, for example failed 3D prints, or common sources of errors and to help users to not only solve their current issue but to also point them towards common vocabulary and concepts, as well as opportunities in the digital and co-located realms to learn more.



#### 4.2.2. *Constructionist “hands-on” didactics*

Both interviewees highlighted the importance of starting to “just make things” and become hands-on with the actual machines and materials as quickly as possible. According to I01, “the only thing that helps [is to bring people] as soon as possible in contact with the machines because the major part is not to teach them something. It is rather to give them a push in the right direction and afterwards be there during the appropriation process. Giving advice, answering questions, and showing some techniques [are important]” (I01). So, constant support from experts whenever something unforeseen happens or when an inexperienced user needs assistance during use is important. Meanwhile, I02 framed his view in this manner: “Make. Hack. Share. Learn. [...] The first step is just to make something. When you have overcome the initial fear of the machine and you have had contact with it, especially when you have been allowed to operate it, [...] then it is quite possible that something gets started simply by doing something.”

Direct engagement is also actively promoted when experts themselves have to perform a task for experience or safety reasons. We observed someone being shown another way of calibrating a 3D printer. This procedure is usually done with a thickness gauge that has to slide between an extruder nozzle and a print bed and has to feel “just right”. So, an experienced user has to do it correctly at least once and let new users “get a feel” for it, which the expert in our observation actively promoted. The general “hands-on” approach in project- and community-oriented settings (see 4.2.3) can be seen to be connected with broader constructionist approaches to learning and sharing expertise (Elliott and Littlefield 1995).

#### 4.2.3. *Expertise sharing within the community*

I01 mentioned different approaches for finding experts or expertise inside the community: “There are cases in which someone contacts people like the other coordinator or me [remotely], as we are easily recognized as experts, and we connect them with each other” (I01). There are also cases in which people merely come by physically and ask, probably because they are already familiar with community-oriented work. Simply walking in during Open Lab sessions is especially easy. Many other possibilities arise, from calling out for help to looking around at what the others are currently doing and seeing if they are available to be asked. Moreover, communication during the Open Lab can be proactively initiated by more experienced members who are interested, by simply asking, “Hey, what are you doing?” or “Can I help you somehow?” (I01). However, for some it is already a hurdle to ask proactively face-to-face for help (I01).

Relatively experienced community members often recognize when someone needs help and consequently offer guidance directly. The interviewees believe that expertise sharing largely works within the Fab Lab in this manner. A defined process is lacking, but when support is required, “someone in the group is quickly asked and then both of them approach the problem together” (I02). This uncomplicated way of requesting expert help and the ad hoc support beyond structured processes has a

positive effect on the learning process (I02). Another approach to getting support from the Fab Lab community is to join its online community boards such as a GitHub repository, and, in particular, its Telegram group (a free, cloud-based instant messaging service), which has a growing number of members (02/ 2020: 170): “What we also have is a Telegram group for the Fab Lab in which you can ask questions, and usually you get a clue or tips quickly [...]” (I02).

However, in some cases, no expert is available within the Fab Lab itself or available for a quick call or chat to try to explain the process of operating the machines or overcoming a problem. Help-seekers may then be obliged to conduct their own investigations, as I02 explained with the following example of a special printer: “I observed it with this Form Labs printer – a special machine with which only a small group of people interacts. There is someone working here mostly on his own because others [are unavailable] to ask for help” (I02).

The Fab Lab was founded around 2013, after both of its current coordinators had first come into contact with 3D-printing and other digital fabrication technologies. Hence, the coordinators’ learning experiences were potentially different from the community-based learning inside the Fab Lab. As a result, we inquired into their personal expertise-gathering processes. Each interviewee had followed a slightly different appropriation process. One had learned more about 3D-printing by himself, with the help of online platforms, such as forums: “[...] to look up things in a forum if a problem arises, I think I’ve still not read the manual. Basically, it’s really forum-knowledge.” The other interviewee had gained his knowledge more through making and trying out things within a small group of interested individuals: “Honestly, I have just tried out everything.” However, in both cases, a community – online and offline – was an essential part of the learning process and its relevance was repeatedly emphasized.

#### 4.2.4. *Interplay of organizational and communal expertise sharing*

Community-based expertise sharing and exchange with other people has been reported as important in co-located settings such as the Fab Lab. However, fixed, organizational points of expertise sharing were also found to exist, namely the safety introductions. These safety introductions are mandatory for users on an annual basis, offered regularly, and shaped by legal requirements such as the need for an actual staff member to deliver them and to have signature-based documentation. The safety introductions focus on the safe operation of the respective tools or machines as well as on rules and regulations. The idea is to “really just show the basics” (I02) and merely offer a brief glimpse into the actual operation and the actual expertise needed to successfully use the machine and the lab in general to their full potential.

Help, co-supervision, and “over-the-shoulder learning” (I01) for this open-ended phase of gaining operational knowledge and expertise are mainly managed through the community. Having encounters with different people with various skill sets and levels of expertise and using the Fab Lab as a space for the community provide the basis for this type of learning environment. Appropriation occurs by “[...] looking at

other's work, then you help each other, then you ask what the others are doing and how it works and that's how it is built up, [...] [this is] the major part of learning inside the Fab Lab" (I01). Emphasis is placed on the role and importance of (non-employed) community members stepping up and adopting the role of experts, allowing over-the-shoulder learning to take place after the initial introduction is completed. With regard to 3D printing, I02 stated that "[...] there are three or four people who are doing this. They have accepted their role as experts and support others with pleasure; they are always ready to help [...]."

Observation: During the slicing process (i.e., the conversion of a 3D model to machine instructions for a 3D printer) of a previously constructed model, one person was unsure about how to read and understand the options and preview the outputs of the software. To find out, he sought help from the other members who were present. He explicitly asked one available person. Then, someone else came and stood beside him and gave recommendations from time to time. The help-seeker operated the computer with the slicing software. Overall, the Fab Lab community seems to have a fluid understanding of expertise ownership and the contextual nature of expertise. Outside of the mandatory safety instructions, in which roles are prescribed, the actual guidance given can change rather quickly, according to need.

Observation: One community member, whom the other members had previously indicated was an expert in printing, was unable to use a specific printer for one of his projects because the printer had a defect. He therefore wanted to use another printer that was currently available. However, this printer was from a different manufacturer and he had never used it before. As the setup of the printer was different from the one that he was used to, he searched for help by asking the other members who were present. He was answered and helped by another community member, who had previously said that the help-seeker was "the" expert for 3D printing.

#### 4.2.5. *Potential for remote support*

Lab coordinators and other community members had already discussed remote connections to the community and other experts, especially to support inexperienced users. One idea was the concept of "a big red help-buzzer next to each machine in the lab that calls a group of voluntary remote experts for that machine who are not currently physically around but who can usually help with operational questions" (field notes). I01 mentioned the benefit of establishing ad-hoc communication with experts to get help or answers, which otherwise would need intensive research and domain-specific knowledge. He provided this example: "The CNC guys (a group of community members who work with professional CNC mills during their day jobs) know which end mill to use for which material, at which velocity. [However,] we have to look into some complicated books of tables every time [...]" (I01).

He also suggested (without this topic being previously mentioned by us) that AR might allow some form of co-presence and perhaps something similar to "over-the-shoulder learning." He also suggested that AR would allow "media richness" during the process. Additionally, however, I01 stated that "over-the-shoulder learning and

getting to make something pretty soon are important parts of the process. I don't think that this can easily be made completely digital and [it] requires a lot of communication" (I01).

Observation: All the interactions transpire around a physical object serving as a shared point of reference. This physical object may be the computer screen during the early stages of construction and slicing, the printer or machine itself, or the finished or failed product of the process. During all of these phases of the production process, the collaborators gather around this object and refer to it, point at it, touch it, and move it during their discussions and explanations. Whether the object was a screen, the extruder of a printer that was being calibrated, or the failed printed artefact, which was used to analyse errors that had occurred during the construction or configuration phase, there was always some physical object around which people gathered.

I02 also highlighted the importance of having multiple perspectives on a problem to effectively solve it, citing as an example the printing bed and the printed geometry, combined with access to information about printer models and settings and some form of face-to-face communication for 3D printing support. He also considered AR as a potential solution and suggested pointing functionality to be able to reference parts of the printer, its output, and the process (see also the visual guides mentioned in 4.2.1).

## **5. Design challenges and implications from the literature and the empirical study**

In this section, we summarize the requirements and design challenges regarding remote troubleshooting and expertise sharing from both the literature (section 2) and the empirical study (section 4) and derive design implications for the support of remote troubleshooting practices. These were used to inform the initial development of a very basic prototype, shARe-it (which is described in section 6). The literature already provided a number of interesting insights regarding remote expertise sharing and ways of supporting it (Table 1). Where relevant, we have also indicated how these insights were reinforced by our own empirical observations.

Design Implication 1: As the literature demonstrates, problems with remote troubleshooting are primarily due to incomplete problem descriptions that are not understood by the expert because they are chiefly based on verbal communication (Castellani et al. 2009; Crabtree et al. 2006; Ludwig et al. 2014). The remote experts must therefore be provided with additional channels to obtain better access to the help-seeker's site and how the problem is presenting itself. The literature suggests focusing on designing a visual communication channel, supported by virtual annotation overlays to enhance the verbal communication (Adcock and Gunn 2015; Fussell et al. 2004; Gauglitz et al. 2012, 2014; Huang and Alem 2011; Oda et al. 2013).

Design Implication 2: Within remote troubleshooting settings, a physical, conceptual, and logical dislocation occurs between the expert and the person seeking

Table 1. Design challenges and implications from the literature

No.	Findings from Literature	Design Challenges	Design Implications
DI1	Troubleshooting often simply relies on an audio channel, and experts have no way of knowing whether their instructions have been executed.	How to enhance the verbal communication through additional channels and provide experts with feedback about a help-seeker's actions?	Provide a visual communication channel to enhance verbal communication and monitor help-seekers.
DI2	The physical, conceptual, and logical dislocation between the expert and help-seeker interferes with remote troubleshooting.	How to enable the expert and the help-seeker to collaboratively use different types of media, so both have access to the same underlying resources?	Provide functionality for sharing spaces and media and using them together collaboratively.
DI3	Gestures support the conversational grounding process. Deictic gestures are efficient for object identification, whereas iconic gestures are useful for procedural statements.	How to ensure that different types of gestures can support the grounding process?	Provide different types of gestural affordances and implement functionality to enable both the expert and the help-seeker to use gestures within their shared space.

help, which increases the difficulty of articulation work (Castellani et al. 2009; Crabtree et al. 2006). A key challenge is how to allow experts and help-seekers to collaboratively use different types of media to support the various forms of articulation work. This suggests a need for shared spaces and ways of providing access to the same underlying resources to foster articulation work (Castellani et al. 2009; Tecchia et al. 2012; Whalen et al. 1988; Whalen and Vinkhuyzen 2000).

Design Implication 3: Gestures play an important role in supporting the conversational grounding process (Clark and Marshall 1981). Deictic gestures are more efficient for object identification (Bauer et al. 1999) and iconic gestures a more effective for procedural statements (Fussell et al. 2004). Remote troubleshooting and the shared spaces must therefore support gestures. Gestures were also seen to play an important role during the empirical study, where they were used to emphasize co-located physical interactions and clearly indicate a referenced object. During the study, the experts used iconic and deictic gestures to support the verbal description of physical tasks.

We can extend upon the three design implications taken from the literature by adding some important design challenges that were derived from the empirical study. Here, the difficult part is to derive implications based on the challenges of co-located expertise sharing that are also relevant for the support of remote expertise sharing. Designing for remote expertise sharing makes a number of the positive aspects of collocation unavailable, such as over-the-shoulder learning, spontaneously noticing

what others are currently doing, and ad hoc questions to other on-site users. Nevertheless, certain design implications for remote settings can still be derived (Table 2).

Design Implication 4: The Fab Lab is a heterogeneous setting with its 300 m<sup>2</sup> and multitude of different machines (CNC milling, 3D printers, soldering stations, etc.) and workstations (manual workstations, semi-automated stations). There are therefore no existing instructions for individual work steps or manuals on how to carry out something. For this reason, expertise sharing often takes place spontaneously and in an ad hoc fashion, with finding experts being heavily based upon community work and simply asking for help. This turns upon being based in the same environment, being able to recognize others who need help, and being able to ask either a specific person known to be an expert or the entire community for help. A challenge within remote troubleshooting settings is to allow these forms of ad-hoc communication and to foster help-giving and demonstration. This implies the provision of a tool that can feed problems to a group of experts, so that an expert can decide if they can and will help. Due to the heterogeneous field and because the occurrence of problems cannot be predicted exactly, the tool will need to take the form of a mobile AR application that allows for hands-free operation of the hardware (e.g. smart glass).

Design Implication 5: As the Fab Lab community is already a field of application in a co-located community inhabiting a complex setting, it already uses tools such as Telegram to communicate, exchange expertise, and provide help outside the physically co-located Fab Lab environment. Within Telegram, the community already posts photos or short videos that support the description of a problem (cp. (Bobrow and Whalen 2002)). As the empirical study revealed, even when people are co-located, it is not always clear in advance who the expert is and which problems

Table 2. Design challenges and implications from the empirical study

No.	Results	Design Challenges	Design Implications
DI4	Finding expertise is achieved through ad-hoc communication and spontaneously asking someone for help.	How to enable ad-hoc communication with experts in remote locations?	Be able to ongoingly feed problems to a group of experts, so that an expert can decide if they can and will help.
DI5	The community maintains a messaging group for organizational announcements and chat-based remote troubleshooting.	How to give access to a pool of experts via already established communication channels, such as Telegram?	Integrate resources such as Telegram to let experts observe the work by offering contextualized information.
DI6	Over-the-shoulder learning is an important expertise-sharing practice.	How to support ad-hoc and visual over-the-shoulder learning from remote locations, where different angles are necessary?	Use a video-feed that is quick and easy to set up and capable of sharing different views.



someone else might be able to provide support for. So, applications like Telegram provide a pool of potential experts that can be drawn upon. Thus, it would seem reasonable to integrate Telegram-like applications in any provision of remote support to act as a communication channel for finding experts. It would also seem sensible to make use of pre-existing applications with which people are familiar, rather than obliging people to learn to use something new.

Design Implication 6: Within primarily hardware-focused co-located settings such as the Fab Lab, experts typically facilitate an over-the-shoulder learning process, in which they examine the machine and things like the printed object together with the help-seeker from different angles and with a shared view. The challenge in this case is how to provide for the ad-hoc and visual nature of over-the-shoulder learning in a remote situation (C6). This would seem to imply a need for a video-feed that is easy to set up and capable of sharing a view from different angles.

## **6. shARe-it: A prototypical AR-based remote troubleshooting application**

Having reviewed the literature to examine how remote expertise sharing is already supported (section 2) and having conducted an empirical study (section 4) that added insights regarding expertise sharing and troubleshooting in co-located, hardware-focused settings, we wanted to build an initial, very rough prototype that could help us to undertake an initial exploration of some of the design implications presented in section 5. Our idea, here, was to engage in some very preliminary low-cost and easy design that could inform the construction of a more concrete prototype once we had established what might or might not work. This initial implementation of the HoloLens-based application, shARe-it, is described below.

### **6.1. Hardware and implementation**

The Microsoft HoloLens' recognition of a physical environment is an important functionality that allows for the spatially aware and consistent placement of digital objects in the real world in combination with HoloLens' projections on a see-through surface. We felt that this capability might extend upon other existing approaches to remote troubleshooting, because other approaches either scan the environment of the help-seeker with 3D-cameras, taking the focus away from the real world (Gao et al. 2016; Tecchia et al. 2012), or scan the environment to provide project guidance Adcock et al. 2013, 2014). The latter is limited in its reach by the fixed position of the projectors. Mobile solutions, such as those proposed by Huang and Alem (2011) or Adcock and Gunn (2015), require additional hardware on the expert's side for capturing their hands or their spatially-projected guides are not consistent. Our initial focus, then, was upon developing as an off-the-shelf solution for a hands-free AR application that could support remote troubleshooting and expertise sharing.

Although we wanted to kick off the design process with a relatively quick and dirty prototype that would be good enough for some preliminary testing, we also wanted to provide the prototype with a solid enough technological foundation to reflect the above motivation. We therefore implemented our application, shARe-it, by using Unity (<https://unity.com/>) and C#. Unity is a cross-platform game engine that can be used to create 3D models and therefore virtual reality as well as augmented reality games. As a 3D game engine, Unity is responsible for rendering the holograms, providing a physics engine and adding to the 3D objects. The behaviour is described by Unity-Scripts, which we programmed in C#. As C# is used by the code provided by Microsoft for the HoloLens, it was used for the entire implementation. We used the Microsoft Visual Studio IDE, which allowed us to deploy the code to the HoloLens itself via USB or WLAN and to use an emulator. We used four main toolkits for implementing the application:

- Microsoft's Mixed Reality Design Lab<sup>1</sup> provides sample app projects that demonstrate how to use various types of common controls and patterns in Mixed Reality. The prefabs and scripts offered by this repository ease the creation of GUIs and establish a consistent visual impression with the rest of the OS, as well as the behaviours.
- Microsoft's HoloToolkit-Unity<sup>2</sup> offers scripts and components intended to accelerate the development of holographic applications targeting the Windows Holographic specifically for use in Unity. Among other things it includes prefabs for setting up input methods, offers interfaces for Air taps, clicks, voice commands or a cursor, as well as a HoloLens camera and assets for adding spatial mapping to a project.
- Microsoft's HoloToolkit<sup>3</sup> adds additional tools for the development of mixed reality applications which are not specifically for use in Unity. Apart from providing the code used in the Holotoolkit-Unity for spatial understanding and to identify surfaces, it contains the server application for sharing data between different mixed reality devices.
- Microsoft's MixedReality Companion Kit<sup>4</sup> contains helpful tools and resources for the development of mixed reality applications. However, these are not intended to run on the HoloLens itself. The majority of the provided tools revolve around sharing the HoloLens' augmented view with someone else using another device.

---

<sup>1</sup> Mixed reality design labs. Retrieved from [https://github.com/Microsoft/MRDesignLabs\\_Unity](https://github.com/Microsoft/MRDesignLabs_Unity)

<sup>2</sup> Holotoolkit-unity. Retrieved from <https://github.com/Microsoft/HoloToolkit-Unity>

<sup>3</sup> Holotoolkit. Retrieved from <https://github.com/Microsoft/HoloToolkit>

<sup>4</sup> MixedRealityCompanionKit. Retrieved from <https://github.com/Microsoft/HoloLensCompanionKit>

## 6.2. Functionality of shARe-it

Within this section, we describe the functionality of the version of shARe-it we deployed. The application was divided into two parts. The first part focused on finding a remote expert, while the second part dealt with the actual remote troubleshooting process (Figure 3). As it was central to realising this first basic deployment, we will focus here upon the role played by the HoloLens.

For our initial design, we conceived of the process as follows: Upon encountering a problem and not knowing how to proceed, a user could start looking for an expert by opening the HoloLens-based AR application, shARe-it. During the initial start-up, shARe-it would initiate an expert finding process by placing a button for finding support directly in the field of view (Figure 4, left). For usability reasons, we enabled an option to move all the elements to the bottom of the view. We also designed the application so that voice commands such as “Find expert,” “Can anybody help me?” “I need help,” or simply “Help” could trigger the process. Although this functionality was implemented from the outset, to make things straightforward, our initial basic exploration of the viability of the system did not look at any of the expert finding elements and focused only on the synchronous expertise sharing.

In brief, prior to any actual synchronized troubleshooting involving direct collaboration between the user and a remote expert, the functionality was designed to work as follows:

As a preliminary, the HoloLens supported the recording of the scene through its built-in cameras and audio recorders, with it being possible to use photos and short video clips to describe the problem at hand (see Section 5, DI1). So, if a help-seeker decided to underpin their problem articulation with visual material, they could initiate this through gestures or voice commands in combination with head movements (Figure 4, middle).

After recording a video or taking a picture (or just capturing an audio file), the current situation could be shared with the community or a specific expert (see Section 5, DI4). In view of its existing use in the community we were working with, Telegram was integrated to make this possible. The user could select the specific type of chat they wanted to use and shARe-it would then offer it to local group, other group chats, or private chats (Figure 5, left). A Telegram bot, BotFather (<https://t.me/botfather>), was implemented to direct and forward the transmitted audio, pictures, or video files, which could handle the communication between the HoloLens

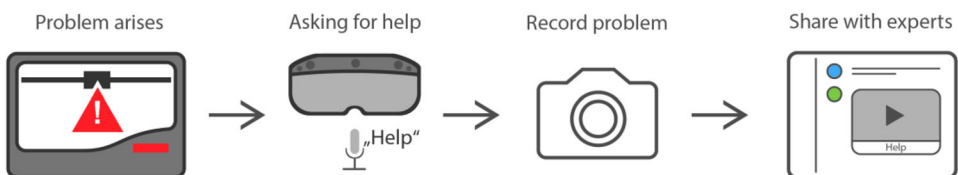


Figure 3. Underlying shARe-it concept



Figure 4. Initial start-up of shARe-it (left); recording a problem (middle); accepting/declining recording (right)

application and the Telegram groups (see Section 5, DI5). After recording the scene, the captured representation could be shared with the selected group, accompanied by a link through which experts could join the troubleshooting session. The link allowed an inline preview of the content to give the experts an initial sense of what might be going on (Figure 5, right).

Once at least one expert had joined the session, the actual synchronized troubleshooting process could start. However, the invitation link remained available to enable other experts to join as well.

As the first step in a troubleshooting session usually involves the help-seeker giving a description to the remote expert (Castellani et al. 2009), the shared inline preview of the content in the Telegram group already encompassed a first partial description. However, we knew that this would not be enough if the problem was more complex or difficult to identify, making it hard for the help-seeker to capture (Lukosch et al. 2015). Prior work has shown that, in these kinds of situations, the users work with the experts to collaboratively narrow down the problem (Castellani et al. 2009). To support this, we therefore implemented a video stream that would enable the expert to observe the situation on their own, in something approaching what might happen in a physically co-located situation (such as the Fab Lab). Our hope was that this would reduce the sense of physical dislocation. Additionally, the HoloLens' cameras was able to provide a shared view that could give the expert the same perspective as the help-seeker, thus potentially easing the communication and reference to objects (Clark and Marshall 1981; Susan R. Fussell et al. 2004).

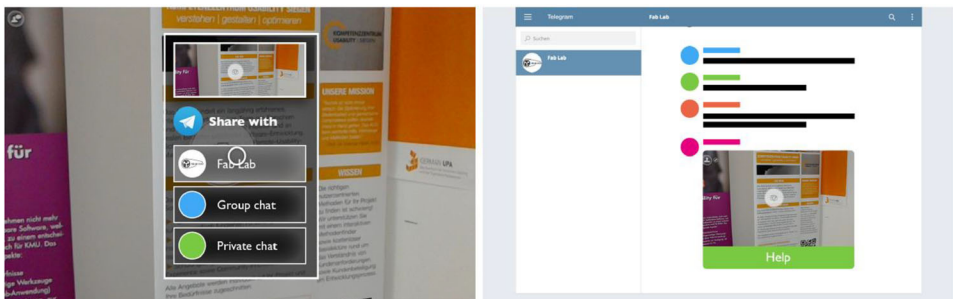


Figure 5. Selecting the target group (left); sharing of the recorded content through Telegram (right)

Even though full view independence for a remote participant can enhance the quality of a collaborative task (Tait and Billingham 2015), we decided to always transfer the “field of view” (camera) of the on-site user, as we wanted to provide an off-the-shelf solution that would support a measure of mobility (see Section 5, DI6). Enabling a full view independence would have demanded setting up a lot more (specialized) kit and we assumed that, through the various channels available, the dislocation could be managed well enough to get the job done. We saw this as a trade-off between practicality, expense and deployability.

During collaboration, a shared understanding of the situation can be strengthened not only by vocal communication, but also by gestures (Fussell et al. 2004). Two of the three tasks in a conversational grounding process are supported by both deictic gestures and representational gestures, which can be either iconic or spatial (Fussell et al. 2004). To allow representational gestures and a feeling of co-presence, we designed shARe-it to not only provide a video from the help-seeker to the remote expert, but to also feature video the other way around. This is already commonplace in remote troubleshooting systems, but we recognised the value of having a representation of the remote expert somewhere within the environment. At the same time, we set things up so that the help-seeker could place the image of the expert where it was most convenient (Figure 6, left). The facial representation of the expert could also be hidden, adjusted, or simply removed. However, we used the hologram to represent an active connection with the remote expert, so, if it was removed, this closed the connection. We made sure this did not close the entire session, as there was a possibility that multiple experts might be in the session or that another expert might join the open session later. When multiple experts were in a session simultaneously, they were each represented by their own hologram with the same respective functionality.

Apart from representational gestures, other gestures can also reference physical objects in a help-seeker’s space (see Section 5, DI3). As both drawing gestures (Fussell et al. 2004) and hand gestures (Kirk and Fraser 2006) have been shown to improve task performance, we focused on supporting both of these (Figure 6, middle). However, we did not implement capturing real hand gestures on the expert’s side to overlay onto the help-seeker’s view. This would have required additional hardware and hampered the ad-hoc communication that we had already found to be essential in the empirical study (see also,

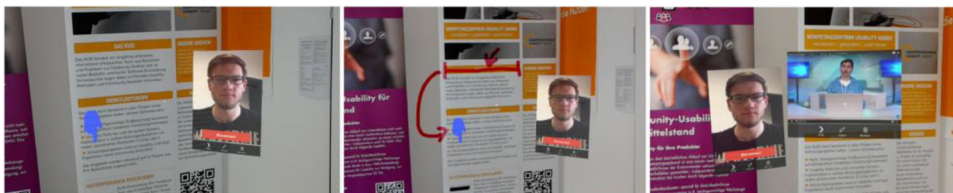


Figure 6. Troubleshooting session with gestures (left), drawings and visualizations (middle), and video integration (right)

Tecchia et al. 2012). We set it so that the gestures would automatically disappear after a brief period so that neither the expert nor the help-seeker would have to remove them manually (Fussell et al. 2004). Given the potential presence of multiple experts, individual gesture markers were distinguished by different colours.

To ensure a mutual understanding of the situation, we also built the prototype so that different types of media could be shared (see Section 5, DI2) (Crabtree et al. 2006). The experts had the option to select various types of media to share with the help-seeker, such as images, documents, websites, and videos. [“There is so much on YouTube, which can really support things on a visual basis.” (I02)]. At the same time, because everything on the side of the help-seeker using the HoloLens was potentially going to be displayed directly inside the expert’s view (Figure 6, right), we felt that the help-seeker had to be able to control the amount of information they wanted to be displayed. We also anticipated that providing all the information simultaneously could obscure some part of the real environment.

## 7. Initial testing

Our main interest in this study was to examine whether new technologies, such as AR-based smart glasses, coupled with new possibilities for shared views and annotations, could offer new possibilities for the support of remote troubleshooting and expertise sharing within hardware-focused settings. Having drawn upon the existing literature and our own situated empirical study to arrive at the initial design outlined above, we wanted to undertake a basic assessment of the viability of the HoloLens-based aspects of the design. The goal here was simply to explore whether our current understanding of what appropriate support might look like and our expectations regarding what the HoloLens could actually provide were sufficient for a fully-developed prototype to be implemented. So, at this point, we were only interested in uncovering evident problems with the technology, before moving towards real-world testing (which is a more expensive exercise in a number of ways).

To make things as simple as possible and to keep costs to a minimum, we concentrated solely upon the functionality designed to support actual synchronous interaction between a user and a remote expert. We therefore developed some simple tests that would include certain kinds of tasks relevant to at least some aspects of a remote troubleshooting scenario. We were also explicit about the proposed scenario for future use with our participants and sought to use the rudimentary tests provided as a prompt to their imaginations by giving them some sense of what modern AR technology could do and what it might be used for. Beyond the tests themselves, then, we cultivated their feedback regarding the system’s possible future use. In this sense, the initial implementation had the flavour of a co-design workshop framed around practical tasks.



## 7.1. Method

The tests involved an initial, pre-interview, a set of group-based tasks using certain aspects of the shARe-it application, and a post-interview. The pre-interview was designed to gain insights into the participants' qualifications and their experience with AR and remote collaboration. The post-interview focused on problems with the use of the application, technical constraints and task difficulties, and suggested features in view of its anticipated future use.

As at this point we simply wanted to examine how AR-based remote troubleshooting and expertise sharing might be best realized, we recruited pairs of test participants, where one would adopt the role of a remote expert and the other the role of a help-seeker. The remote expert used a notebook or smartphone and the help-seeker used the HoloLens. In total, we undertook 7 sets of tests with 14 participants (P01-P14) grouped into seven groups (G01–G07). The participants were recruited at our university and within the Fab Lab. Most of the participants were students or research assistants and a few already had experience with AR and the HoloLens (see Table 3).

The tasks we set the participants sought to simulate certain aspects of remote troubleshooting, with a help-giver (the 'expert') and a help-seeker attempting to identify a problem and then to try to solve it (collaboratively) using shARe-it. The expert therefore had to guide the help-seeker through different physical activities. To keep the implementation as simple as possible, we placed both the expert and the help-seeker in one room, but kept them separated by placing a wall of cardboard boxes in the middle of a table and sitting them at opposite ends of the table (see Figure 7). The boxes were sufficiently high that the participants could not see over them to the other side and thus had to rely on the shared view provided by shARe-it. At the same time, they could still see and talk to each other, as would be the case with the face representation and audio channel in shARe-it. We therefore initially switched off the face representation for the tests, but it turned out later that this functionality was switched back on by one group and used for media sharing (see section 8).

We focused on two different tasks. The first was a LEGO assembly task (cp. (Gao et al. 2016; Huang and Alem 2011; Kirk and Fraser 2006; Tecchia et al. 2012)). Using a LEGO assembly task (Figure 8) demanded little in the way of set-up of other equipment outside of shARe-it. In addition, the task required various kinds of physical manipulation that had some relationship to the tasks encountered when troubleshooting. Of course, the help-seekers had some understanding of the artefact that would not necessarily be the case with a complex machine, but it still gave us an opportunity to understand what resources were called upon for guided manipulation and whether these resources were being readily found in the shARe-it application.

We gave the experts instructions regarding the building of LEGO artefacts and the help-seekers had to build them. The expert had to guide the help-seeker through the process of systematically building the artefacts by using the various features

Table 3. Test participants

No.	Role / Job	E d u c a t i o n a l Background	Experience of using HoloLens
P 0 1 (G01)	Professor	PhD information systems	Supervised student work about AR supported hardware tutorials
P 0 2 (G01)	Research associate	Master's degree information systems	Exploring VR possibilities in a research project
P 0 3 (G02)	Research associate	Master's degree media business	Worked with the HoloLens during a project
P 0 4 (G02)	Research associate	Master's degree mechanical engineering	AR and HoloLens use in research project, externally developed
P 0 5 (G03)	Master's student HCI	Bachelor degree electrical engineering	No previous use of AR, use of VR for 360° pictures and videos
P 0 6 (G03)	Master's student HCI	Bachelor psychology	Experienced with VR/AR, managing a VR lab. Mostly focused on VR
P 0 7 (G04)	Bachelor student information systems	High school	No previous use of AR or VR
P 0 8 (G04)	Bachelor student information systems	High school	No previous use of AR or VR
P 0 9 (G05)	Electrician	Vocational training	No previous use of AR or VR
P 1 0 (G05)	Nutritionist	Bachelor degree nutrition science	No previous use of AR or VR
P 1 1 (G06)	Industrial management assistant	Vocational training	No previous use of AR or VR
P 1 2 (G06)	Bank clerk	Vocational training	No previous use of AR or VR
P 1 3 (G07)	Apprentice in metal industry	High school	No previous use of AR or VR
P 1 4 (G07)	Machine Operator	Secondary school	No previous use of AR or VR

available in shARe-it. Although this might seem to have an experimental motivation, our interest was less in measuring performance than in trying to unpick how the interaction between the 'expert' and 'help-seeker' proceeded as an interactional, collaborative and negotiated exercise, given the support of the technology. To



Figure 7. Test setting

identify the benefits of the application, we used two slightly different variations of the task. In the first session, the participants only had the shared view as a supporting technology; in the second session, the gesture and drawing functionalities were enabled.

Without going into detail at this point, it should be noted that the setup of the first test had some limitations because the physical LEGO objects were small and the participants encountered difficulties with the precise identification of the physical surroundings caused by the HoloLens spatial mapping (see section 9). The second test therefore focused on larger physical Lego-based objects that would fit better with the HoloLens's mapping (Figure 9). Here, the objects had to be placed in a specific position and rotated (cp. (Adcock et al. 2013; Adcock and Gunn 2015)). To provide a basis for the communication of the positions of the objects, they were situated on a  $5 \times 5$  grid. We used five different objects, each one unique in terms of colour and shape (Figure 9). Two of the objects had a similar shape, whereas two had the same colour, such that no object could be solely identified on the basis of a single attribute.

The first test indicated that the shared view was more beneficial than the marker and drawing options. Therefore, we differentiated between three types of options: (1) without the HoloLens and just using audio; (2) with the HoloLens and shared view enabled; and (3) with the HoloLens, shared view, and markers and drawings enabled. All participants undertook all tasks across both tests.

Prior to the actual tests, we introduced the overall handling of the HoloLens, the application itself, and the different types of tasks. In order to learn the general handling of the HoloLens and to calibrate the tool for the respective participants, the participants were asked to enact the first phase, “finding an expert”, even though the backend to support this was not enabled. For this purpose, the participants were asked to record a video in which they explained the current situation by using gestures or voice commands and then sent it via Telegram to an account predefined by us using the air tap gesture. This gave us some basic insights regarding the current “finding an expert” process and gave the participants some training in the use of the application. None of the participants had problems with handling the technology at this point, so we will focus here on the collaboration results and subsequent suggestions.

In view of our overall interest in understanding the viability of new AR technologies for supporting troubleshooting, a key focus of the tests was the part played by the usability of the technology. During the tests, the participants were therefore asked

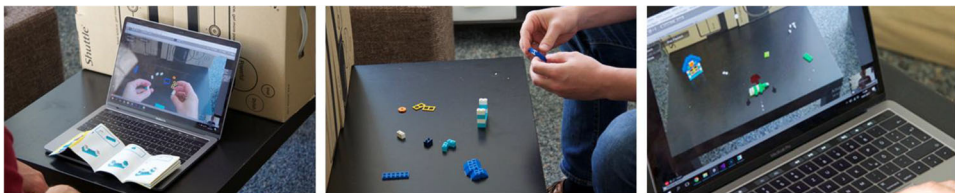


Figure 8. LEGO assembly task



Figure 9. Positioning and rotating task

to “think aloud” (Nielsen 1993) and were audio-recorded. The tasks themselves were recorded using screencast software, which recorded the screen of the remote expert and the shared view of the help-seeker wearing the HoloLens. In addition, the webcam on the laptop or the front camera of the smartphone recorded the expert to capture any gestures they performed via the face representation channel.

## 7.2. Results

The tests were conducted in German. All of the excerpts in the following were therefore translated from German into English. We have tried to capture as closely as possible the sense of what was originally said.

### 7.2.1. *Establishing a common ground*

To successfully communicate the type of object a help-seeker should use (object identification) and the form of activity to undertake with the object (procedural statement), the remote expert and the help-seeker had to find a common ground of understanding. The shared view of shARE-it made arriving at a common understanding easier. After using shARE-it with the sharing enabled, an expert (P05) said about the first type of task, where it had only been supported through audio: “If I had seen it before, the step until the other person understood what I meant would have gone much faster. Then, P06 could have placed the brick there and I could have said ‘No, that is wrong. Turn that a little bit further,’ or something like that” (P05). The benefit of grounding the process through the shared view was more obvious in the groups with the object placement task, which included an additional step without the shared view. This resulted in multiple situations in which both participants had to find a shared understanding of different parts of the object and hence communicate its alignment, as illustrated in the following vignette:

P05: The pike points to the top right.

P06: I’m not quite sure about the “pike to the top right.”

P05: You have something like an “L,” and the pike of that is on the top right.

(P05 draws an L shape in the air, stopping at the point that he refers to as “pike.”)

P06: But what do you mean by “pike,” this part?

(P06 holds the object up, so that P05 can see it over the cardboard boxes, and points at one end.)

P05: No.

P06: That one?

P05: No. The other one.

P06: That one?

P05: Yes.

P06: The corner?

P05: Yes. Ok, then we will call it the “corner.”

In this situation, an equivalent of the face representation channel was used to overcome the misunderstanding between the participants. As a purely verbal description of the problem was not sufficient, the participants in this group simply violated the setting and showed themselves Lego bricks over the “wall”. On the one hand, this shows that the setting was not as well-constructed as we might have hoped. On the other hand, however, it shows the urgent need for visual support, since purely verbal descriptions quickly reach their limits. [“If you can see and hear the other person, you can be substantially compensated by using other ways of communicating.” (P02)] The face representation channel therefore supports the process of finding a common ground by adding the scope for things like visual clarifications (which turn in their own right, upon various forms of deixis, e.g. ‘That one?’, ‘No. The other one’) as can be seen in the way this group used it to show the object to each other to make sure they both meant the same feature (Figure 10).

### 7.2.2. *Communicating object identification*

One aspect of establishing a common ground was the identification of the object or the LEGO brick. During the tasks without digital markers, the identification was only communicated verbally. To identify the different objects, the participants opted to use either previously established descriptions as a reference, make up a new one when it was needed, or use common knowledge, such as the shapes of different letters (T or L), colours, or a combination of the above. As the two excerpts from different groups below show:

P01: Next is the “L” which is on the top-left for you.

P03: Then you’ve got a red brick, similar to a” T.”

P05: The “Tetris” thing is in the first row from the top, in the middle.

P05: Take that brick and push it there.

P05: I’ve marked one brick and that is the target.

On the one hand, the shared view did not change the verbal identification process. The observable benefits of the shared view included the improved capacity to support the procedural statements and its use as a feedback channel. On the other hand, the digital markers, augmenting the shared view, did alter the communication between the participants without the interference of the constraints of shARe-it and the HoloLens hardware: “Then you have to communicate which brick you’ve meant. So, you can either say ‘Brick xy,’ or you can do it like I did it in the last step, clicking on the brick and placing a marker on it” (P05).

The group that was the most experienced with AR maximized the use of the digital markers. However, most groups encountered difficulties with the markers’ accuracy during their session (this prompted us to change the tasks from the LEGO assembly to the object placement task). G01 used the feedback provided by the shared view to support the identification process. The expert started to describe the LEGO brick needed for the next step, while the help-seeker searched for it and pointed at the brick he thought matched the description.

Where the selection was correct, the expert confirmed it. Another reason for the less frequent use of digital markers during the object identification might have been that the identification of the different objects was not so much of an issue as it was with their assembly during the first task. Often, if small uncertainties occurred, they were solved beforehand.

P01: Next is the thing that lays on the top-left on your side; exactly this.

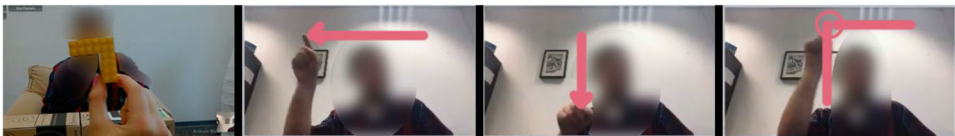


Figure 10. Help-seeker showing object to expert (left); expert showing what he means with “row” and “column” (right)



(P02 points with his index finger at a LEGO brick.)

P01: Some turquoise-colored, big, exactly that one.

(P02 grabs a brick during the description.)

### 7.2.3. *Communicating procedural statements*

The next important part to guide the help-seeker through her task is to specify the action to be taken with a previously identified object. In terms of the LEGO assembly task, the action involved defining where the next brick had to be placed to complete the construction of the LEGO model. For the object placement task, the pertinent position of the object was inside the grid and its alignment. As with the object identification phase, this task was mostly performed verbally without using digital markers.

During the object placement task, the position of the object was communicated through the grid system. One participant suggested using a chessboard-like system to uniquely identify the different fields [“The rows are letters and the columns are numbers” (P06)], which amounts to explicitly creating a shared frame of reference. Most groups adopted a similar approach that used the number of a row and a column in combination with an origin from which to count. During the LEGO assembly task, the group had the shared view at the outset and therefore used this shared view as a feedback channel to help them to narrow down the instructions.

P03: That should go to the left in the uppermost row.

P05: In the first row, in the second column, counted from left.

P05: The” Tetris” thing is in the first row in the middle.

P09: Each of them has to be on the outside edges, exactly.

A method used increasingly with the introduction of the shared view was the use of other objects as points of reference. This mode helped the participants to communicate the positions of the objects more rapidly and easily. Most groups performing the object placement task started to use a sequential approach at some point to order the instructions for the different objects. They would either go from the top row to the bottom or the other way around, adding the information about which object and which column or the lack of an object in that row.

P01: Where the dark-blue one was, there you put the white ones.

P03: The red one, in the second field from the left, in the same row the yellow one is in.

P03: Next the yellow 'L' in the field diagonally below the red one.

P03: Let's start with the uppermost row, there we have the yellow 'L'.

P13: Now we will go from the bottom to the top.

Within the object placement task, the digital markers were used more than during the assembly task as the task setup was more customized to the constraints of the prototype and hardware: "You simply don't have to say so much anymore, you can simply place a marker there and the other person knows, that's the target" (P05). However, some groups had issues with using the markers for visualizing the placements due to the constant movement of the other person's hands and they therefore used them less and went back to plain verbal descriptions. During the use of the digital markers, the verbal communication was largely reduced to expressions such as "there," "in this field," and "where the marker is."

Aside from the position, the expert also communicated the alignment of the objects to the help-seeker. Given the lack of a grid system for the positioning, the participants had to rely on their shared understanding of the objects' parts to convey their rotation. Reaching that shared understanding and creating a common ground for the communication could take some time and back-and-forth interactions between the group members. Further referential expressions such as "horizontal" and "vertical" were used. At this point, additional descriptions were necessary regarding the part of the object that should be aligned in a given way. The intensified complexity of the information to be transferred increased the emergence of questions to seek confirmation that the information had been understood correctly.

During the LEGO assembly task, the face representation channel was used not only for the alignment but also for the transfer of other information. P05 also came up with the idea of using the face representation channel for sharing media with the help-seeker in the form of the instructions' paper. This was an intended possibility, as media-sharing capabilities would increase the usability of the process by providing a channel dedicated to the sharing of information in the form of media. P03 attempted to demonstrate the alignment of the object via this option and encountered some problems. ["As I showed it, I had to keep in mind that it was mirror-inverted." (P03)] Meanwhile, P05 used the channel to point in the direction he was currently mentioning, only to also encounter problems with the mirror inversion. The introduction of the shared view eased this process. As P03 put it: "If you see it through his eyes, it makes it a lot easier. You can simply say 'rotate it 90 to the left.'" This behaviour could also be observed by using the shared view as a feedback channel and allowing the experts to give procedural statements iteratively instead of trying to describe the

full action at once, which reduced the effort of verbally expressing the instructions (e.g. P03: “now take the other piece” and “now put that next to the previous one”).

#### 7.2.4. *Communicating task status*

Monitoring the task status is one of the most important aspects during remote troubleshooting (Fussell et al. 2004). However, without a shared view, the expert lacked visual feedback and had to rely on the verbal feedback from the help-seeker. As one expert noted: “During the first step, I had no feedback at all. I had to trust that you had done it correct, but I never really knew if you had done it right” (P11). Conversely, the help-seeker stated: “I hope I’ve done it right” (P012). The expert had to rely on the help-seeker’s understanding of the verbal descriptions and that the help-seeker would perform it without error. As the expert had no feedback, she could not locate the mistakes the help-seeker had made and therefore could not correct them. As one group had previously come up with the idea of using the face representation channel for media sharing, they also used it as a feedback channel by showing the instructions to the help-seeker after they thought they had completed the task to check if they had made any mistakes.

The participants used the shared view to iteratively give instructions. The shared view also allowed for monitoring the task status, intervention, and correction when mistakes were noticed or a mistake was made in following the instructions (Figure 11):

P01: The little one has to go right there on the top, there on top of the blue one.

P02: On top of this one?

(P02 points at brick.)

P01: Yes. Listen.

(P02 places brick at the position.)

P01: No, just a little bit higher. That’s right.

Although the digital markers were also shared and could therefore be used for communication, they were seldom used for feedback. As discussed in the section about procedural statements, the markers were used for giving instructions based on the shared view. The fact that the markers were also visible to the expert gave them feedback about the success of their actions. This was especially helpful in the case of uncertainty, as the participants could sort out inaccuracies by providing feedback. Sometimes the help-seekers also placed markers to confirm the position they had

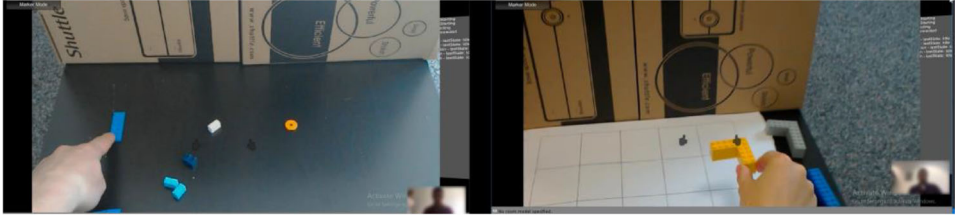


Figure 11. Help-seeker pointing at a brick he thought the expert was referring to (left); monitoring the execution (right)

understood, which was sometimes needed because of problems with spatial mapping (see 7.2.5).

#### 7.2.5. *Prototype and hardware constraints*

One issue was that the field of view provided by the HoloLens cameras was not exactly the same as that of the help-seeker. This resulted in the experts having to guide help-seekers to steer the view toward objects they wanted to see:

P01: You are a little bit out of the viewport, could you... ah thanks... that's better.

P06: Can you see that when I look at it like that, or is it outside the viewport?

P05: No, I cannot see it. Yes, that is better.

P05: Could you please look up a bit.

The field of view was limited not only by the viewing angle but also by the near clipping range, which is the minimum distance objects have to be away from the HoloLens to be rendered. This aspect caused confusion, with markers disappearing without any obvious reason, because the HoloLens provides no feedback for this phenomenon. ["There is no feedback that the distance is too low at all." (P03)] Within the assembly task, the groups struggled the most with these limitations because the task demanded that the help-seeker build a LEGO artefact, so they often had to hold it in their hands. The groups often countered the problem by trying to hold the artefact as far away as possible, but an arm's reach was not always enough (Figure 12).

The biggest problem of all was the unreliability of the HoloLens spatial mapping, which is the basis of the spatially-aware placement of the digital markers. During the LEGO assembly task, the precision of the markers was so poor that the group barely used the markers at all and more or less resorted to using just verbal communication, completely negating the purpose of the exercise: "From my perspective, it seemed [...] that the surface of the table was not

recognized correctly. It was all about 10 centimeters below the surface; therefore, the markers all pointed to nothing” (P02).

P03: Oh, that is really shifted away from its intended position.

P04: I would put it here now.

P03: I should go on field to the left.

P11: Where did the marker go now?

After initial completion of the task, one group tried a similar task again, but this time placing the grid on the floor instead of the table. This approach eliminated the surface problems and allowed the communication to be reduced to a minimum, as the markers could be successfully used to identify objects and their destination, only needing verbal descriptions of their alignment.

Although the markers were helpful, at least for the communication of information describing an object’s position, the participants still mentioned that the shared view was much better than the markers: “For me, the improvement was not the pointing at all. Instead, it was having the view that P02 has” (P01). One issue pertained to the alignment of the marker and drawings itself. The implementation guarantees that the marker always faces the user at the time of the creation but only on the Y-axis. As the tests revealed, this is insufficient for real use scenarios because the help-seeker may move around, upsetting the visibility and coherence of a marker’s direction (cp. (O’Neill et al. 2005)).

#### 7.2.6. Conceptual improvements

Aside from critiques of the usability of the HoloLens and shARe-it, the participants had also suggestions for additional functionality. A functionality they missed was the option to pause or freeze the video to better mark a point, as the head movement of the help-seeker could cause difficulty in precisely specifying a point: “For me, the use of the markers was hard when P06 moved her head a lot. I think Skype has something

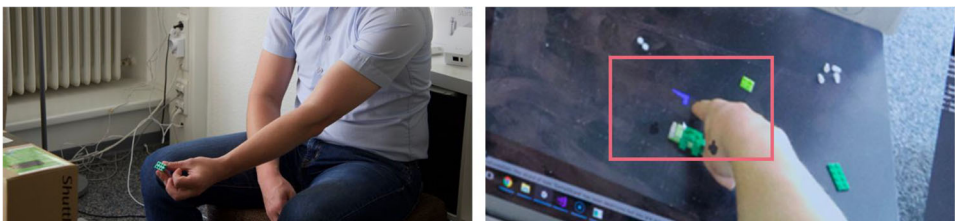


Figure 12. Holding the artefact at arm’s reach to avoid the near clipping range (left); marker cut by the near clipping range (right)

where you can freeze the image and then draw or place a marker or something else. That [feature] would be helpful” (P05). This statement came from one of the notional remote experts. Whether freezing the video would actually have a positive effect on collaboration remains an open research question (cp. (Fakourfar et al. 2016; Tait and Billinghamurst 2015)).

Another suggestion was to increase the discernibility of the sequence of instructions. When the experts had to give instructions in a specific order (as in the LEGO assembly task) or had to first identify an object and then define its position (object placement task), a means of clearly distinguishing the steps could be helpful. This distinction could be expanded by introducing different symbols to further specify the tasks (e.g. rotating by a specific degree).

The participants also suggested the use of a touch-based approach that could digitally remodel the physical objects and allow the remote expert to place them correctly within the help-seeker’s view. This functionality combines the position and rotation into a simple drag-and-drop interaction. However, such functionality requires the scanning of the environment to be sufficiently precise to identify the objects and their properties and render them as manipulable objects for the expert, which is not (yet) possible with the current version of HoloLens. This is particularly the case when dealing with objects or machines that are not necessarily known in advance and therefore cannot be prepared in advance (unlike the situation with Castellani et al. (2009), for instance).

## 8. Discussion

In 2009, Castellani et al. (2009) implemented a screen on a printer that captured the exact state of a printer and which was then shared between both parties, the troubleshooter and the help-seeker. Here, the expert could then manipulate features of the shared representation and draw arrows or annotations on it. Based on their technology, they already envisioned AR technologies for supporting remote troubleshooting to counteract the various dislocations that arise due to remote settings.

Nowadays, many AR-based approaches promise to support remote troubleshooting and remote expertise sharing (Adcock et al. 2013, 2014; Fakourfar et al. 2016; Gao et al. 2016; Gauglitz et al. 2012, 2014; Gurevich et al. 2012; Huang and Alem 2011; Poelman et al. 2012; Tait and Billinghamurst 2015; Tecchia et al. 2012). Our initial assessment of the potential of the shARe-it seems to suggest that modern AR technologies, such as smart glasses can further improve the remote troubleshooting process. In contrast to existing approaches, our starting point in this study was to examine co-located settings of expertise sharing and troubleshooting to identify the relevant practices and then to reflect upon what aspects of these practices might be effectively realized in remote troubleshooting settings. Most studies tend to look directly at remote troubleshooting settings and consider how to support existing remote practices. In our case, we tried to set aside any pre-existing assumptions about remote practices.



One of the big differences between situated co-located assistance and remote troubleshooting and expertise sharing is that the participants no longer have all of the grounds necessary for intersubjective interaction available to them (cp. the concept of ‘fractured ecologies’ first discussed in CSCW by Luff et al. (2003)). Most existing AR approaches try to support remote troubleshooting through new types of interaction such as markers or drawings. However, as even our relatively basic tests revealed, it is not necessarily the AR-specific functionalities of such smart glasses (e.g. marker) that provide the decisive factor in supporting troubleshooting. The major benefit of these technologies is the shared view and especially its resemblance to the help-seeker’s view – even if the visualization is not completely accurate.

The shared view tries to provide a common ground and a necessary feedback channel. Being able to monitor the task status allows the expert to instantly comment on the actions of the help-seeker and correct them if necessary. The digital markers then become, in best case scenarios, helpful as an adjunct to this, assuming that they can be placed accurately enough. However, this could not always be assured and even if it worked, it did not cover all expectations. The right position could only be established iteratively. During the pre-study, the experts did not have to engage in iterative verbal communication to guide the help-seeker to the point of reference. During the tests, however, iterative deixis was a commonplace strategy used by the non-co-located actors to overcome fractured ecologies where a mutually available intersubjectivity could no longer be taken for granted. Right now, then, AR-based technologies do not do enough to repair the loss of intersubjectivity.

This emphasizes an ongoing weakness of AR-based approaches. Verbal descriptions are often more precise within such collaborative negotiation processes than simply pointing to individual physical things. The digital markers serve only to focus the view on a specific area, but the decisive factor remains the language that describes the respective object. Even though many approaches focus primarily on supporting visual input options such as freezing videos (Kim et al. 2018) or stabilizing the digital markers (Fakourfar et al. 2016; Gauglitz et al. 2014), the participants in the tests were vocal in their requests for features that might better describe and communicate an objects’ alignment. Approaches should therefore explore further support of the actual communication process. Examples could include physical objects being recognized via machine learning methods and computer vision (Gauglitz et al. 2012) and, when focused upon, given a label in order to establish a shared terminology between those seeking help and remote experts. Another possibility is that, when thinking about complex machines with a multitude of parameters, the additional visualization of sensor values or machine states could support a remote expert.

Shared views and visual support are not only feasible using AR. Instead, any setup with a camera and suitable streaming components could support the troubleshooting process. Even the ego perspective is not unique to the HoloLens; a head-mounted camera can achieve the same effect. However, the HoloLens offers these capabilities in a ready-to-use single package, requiring little or no additional effort and it does not hinder the help-seeker’s mobility. Nevertheless, the HoloLens currently still has a

number of challenges to overcome. Even if we could have achieved similar results with a non-AR device, a visual communication channel's capacity to significantly enhance expertise sharing would still need to be confirmed. The shared view functions as a shared virtual space that all of the participants can use and refer to, reducing the physical and conceptual dislocation and enhancing the scope for unambiguous communication (Castellani et al. 2009). The help-seeker's view supports not only the visual aspect of over-the-shoulder learning but also a feedback channel for the expert to monitor the help-seeker's execution of an activity. The literature, the empirical study, and the initial tests of our HoloLens-based design generally agree that the shared view is the most important aspect of enabling efficient remote troubleshooting through AR.

Our application, shARe-it, combined the advantages of the ego perspective of the help-seeker with an accessible and user-friendly hardware setup. In comparison to other projects (Adcock et al. 2013, 2014; Gao et al. 2016; Tecchia et al. 2012), the HoloLens reduces the effort needed for the help-seeker to share their view. The HoloLens provides a shared view and a capacity to observe a problem from different angles; furthermore, it has the requisite built-in sensors for scanning the environment. The capability to scan the environment is essential for allowing remote experts to reference objects in the help-seeker's environment. Coupling this capability with the functionality of being able to bring several experts into the interaction spontaneously fosters the ad-hoc expertise sharing that we observed during our co-located empirical study.

Gestures were used to identify objects, enrich verbal descriptions of actions, demonstrate movements, and/or show distances or positions (Bauer et al. 1999; Fussell et al. 2004). Fussell et al. (2004) identified drawing gestures to be the most efficient, Kirk and Fraser (2006) favoured natural hand gestures, and Kim et al. (2019) suggested a combination of sketches and hand gestures. The initial shARe-it design focused on both drawing and pointing gestures to make "full use of both pointing and representational gestures to ground [the collaborators'] conversations" (Fussell et al. 2004). Although augmentation through digital markers to provide pointing gestures exhibited potential for simple tasks such as the ones used during the tests, however, the usefulness of these markers suffered considerably from technological constraints, affecting the accuracy of the markers and therefore reducing their use.

Although these results could demonstrate the usefulness of digital markers, they also clarified that digital markers are only beneficial for object identification and simple procedural statements. For the most part, object identification was at least as simple using only verbal communication. This result might be explained by the simple and uniquely identifiable nature of the objects used in the tests. In the case of objects that are more difficult to distinguish from each other, the markers might be more helpful. However, the procedural statements already showed the limitations of this approach, as it does not offer adequate support for describing complex tasks that require substantial articulation work.

Nonetheless, when the accuracy is sufficiently high, the markers may significantly reduce the amount of verbal communication needed to perform tasks. In this case, the description of the desired position could be reduced to a simple phrase such as “there.” Some of the technical problems occurred because of the HoloLens’ hardware constraints and the prototype’s implementation could be tackled in a different manner (e.g., via Skype for HoloLens). Skype for HoloLens uses standard video transmission. Additionally, it allows for setting markers, drawing and the sending of images. Setting markers and drawing inside the spatial surroundings of the HoloLens wearer can be handled by freezing the view, taking a photo of the moment, adding the augmentations, and then returning to the live view. This is possible because every photo taken by the HoloLens contains information about the virtual camera object and its relation to the physical surroundings, therefore allowing the calculation of the 3D coordinates of specific points in the photo. This aspect solves the problem of the HoloLens wearer having to keep her head steady while the remote user adds the augmentations. However, the concept used in the initial design aimed to offer a live view of the help-seeker, not only to enable constant feedback, but also to mimic the co-located process as closely as possible. Having undertaken the tests, it is clear that this approach has to be reconsidered and an optional function to freeze the view might combine the advantages of both approaches. As this research indicates, a viable approach might well involve automatic freezing instead of the manual freezing of the frame (Kim et al. 2018).

Sharing media was an important part of the concept. Although we did not focus on media sharing during the tests, one group used the face representation channel and attempted to share the instruction paper given to the expert. This example illustrates that sharing media simplifies some aspects of remote troubleshooting, as the sharing of instructions would ease the entire articulation work. Instead, the expert would only have to share information with the help-seeker that she did not have or could not find on her own. Referring to documents and using them to provide the information needed is a role that an expert normally performs (Yamauchi et al. 2003). Here, it is important to ensure that the information should match the knowledge of the help-seeker. As Castellani et al. (2009) showed during their re-design of the Xerox knowledge-base, which aimed to support remote discussions with help-seekers, the rather technical language used in typical knowledge bases can often be virtually unintelligible to help-seekers. The information an expert is using tends to be formulated upon an assumption of a certain degree of help-seeker knowledge. AR technologies make it possible to show instructions in a simple way that can mediate between expert knowledge and help-seeker understanding (cp. Hoffmann et al. (2019)).

For an optimal wearing position, the HoloLens needs to get most of the expert’s screen inside the user’s view. This is not always easy or comfortable. The holographs are limited, as they can never fully cover the background and treat black as opacity or an alpha channel, which can often preclude against the use of black. As a result, reading can be difficult if the environmental settings are not perfect, which is usually

the case in everyday situations. The constraints added by the see-through aspect of the HoloLens display did not occur in a setting using a traditional monitor (Gao et al. 2016; Tecchia et al. 2012), but they forced the help-seeker to look away from the work she was doing. In addition, when not using a controlled environment (Adcock et al. 2013, 2014), these kinds of systems can suffer from environmental lighting conditions, or other contextual factors that occur during practical everyday use. To sum up, then, despite its potential for remote troubleshooting, the tests revealed that existing off-the-shelf AR-technology such as the HoloLens has some way to go before it can provide for forms of interaction anywhere near as seamless as the interaction visible in co-present help-giving situations.

## 9. Conclusion

During remote troubleshooting, numerous dislocations complicate expertise sharing between a remote troubleshooter and a help-seeker (Castellani et al. 2009). Supporting remote troubleshooting therefore requires the establishment of a shared understanding and the presence of shared objects that the collaborators can relate to and use to focus their interactions around, as they would in a co-located situation (Crabtree et al. 2006). Augmented reality enhances the features of the real world with multi-modal virtual information (Azuma 1997). AR also has the potential to address different types of dislocation and foster a shared understanding by enriching verbal communication through a visual channel (Castellani et al. 2009), which encompasses gestures and new ways of identifying objects, and thus supports conversational grounding (Bauer et al. 1999; Clark and Marshall 1981). Many approaches already try to support troubleshooting and expertise sharing in remote situations.

Based on the state of the art, we have illustrated the current challenges that occur when supporting remote troubleshooting with AR. The video functionality, in particular, tries to support the essential aspects of troubleshooting. However, as the preliminary testing of our AR-based application, shARe-it, has shown, functionalities such as markers and drawings only offer limited further support. If current AR technologies demonstrate certain weaknesses even under the simple conditions established for our tests, these weaknesses are likely to be intensified in more complex settings with more complicated processes, such as those in a Fab Lab. Lukosch et al. (2015) argue that solving these types of problems will “still require a team of experts to physically meet and interact with each other.” This case is especially true when identifying the problem itself is perceived to be an issue (Ludwig et al. 2014; Piirainen et al. 2012).

Drawing upon an empirical pre-study in a Fab Lab, we revealed how co-located troubleshooting and expertise sharing is accomplished and the factors that play an important role. Based on this pre-study and extant findings in the literature, we derived design challenges and implications for remote settings, which led to our initial design of the HoloLens application, shARe-it. shARe-it comprises the ad-hoc finding of experts from an expert pool, based on chat technologies, and a subsequent,

AR-supported process of problem identification and solution. However, when we undertook some basic testing of some of the synchronous aspects of shARe-it, we uncovered a number of challenges that will need to be overcome to support remote troubleshooting through AR. These challenges primarily relate to the capacities of the current hardware (with the Microsoft HoloLens being one of the most advanced AR technologies available), as well as the appropriate troubleshooting process support. Overall, this design case study has served to underscore the current obstacles confronting AR technology and the need for further research to address the complexity of the communication support, beyond simple tasks in simple settings and into real world practice. With this paper, we contribute not only to CSCW by providing an empirical study regarding how co-located expertise sharing takes place, but also to an understanding of the current challenges and remaining false assumptions that can be seen to be associated with using modern AR technologies to support remote troubleshooting. In particular, we have highlighted the need for designers and developers to primarily focus on supporting both verbal communication and, beyond this, the provision of effective resources for intersubjective reasoning in order to mitigate dislocations within remote troubleshooting.

Our study has some limitations. First of all, in the empirical pre-study we only interviewed a small number of experts, which makes our dataset seem small. However, the empirical study in this paper complements the expert view regarding the challenges faced by end users in dealing with complex technologies that has already been published in Ludwig et al. (2014). As there were only two Fab Lab coordinators at the time of the empirical study, we could not call upon any other experts with such extensive knowledge. Secondly, we only conducted some basic, very constrained testing of our application and have yet to properly evaluate all of its functionality in proper troubleshooting settings. The findings are therefore limited to the issues confronting the use of AR to mediate practical collaborative tasks involving guidance and instruction. However, we are currently planning to roll out a revised and developed version of shARe-it that draws upon these initial findings. Evaluation of this new version should deliver far more detailed feedback about the use of the application in actual troubleshooting practice. We are also currently working with rescue workers to explore possible use cases for shARe-it in the support of in the field first aid provision and to establish new communication channels between first responders and control centres. A third limitation is that, as the participants in the tests were co-located, we were not able to fully assess the effectiveness of the face representation channel, though we could already see that the face representation channel would be used in other ways than we had originally intended (Orlikowski and Hofman 1996).

Apart from rolling-out shARe-it within the Fab Lab, our next steps include combining it with sensor data gathered from machines (Ludwig et al. 2017) or from 3D models (Jasche and Ludwig 2020) and providing experts with this environmental and contextual data to help with the identification of problems (Billinghurst 2013). We will also be trying to integrate computer vision functionality to identify and classify physical objects so as to overcome terminology issues. In addition, we will ensure in our next version of

shARe-it that experts are able to directly share their experiences through shARe-it with the community via Telegram (without a prior help request). In this way, other members can observe their practices and learn from them, as is currently the case in co-located settings through over-the-shoulder learning.

One further, open research question that emerged from our empirical pre-study was how the co-located noticing of people seeking help and the proactive addressing of problems (“Can I help you somehow?” (I01)) by the experts can be transferred to remote settings. What this implies is the remote noticing of troubles and a capacity to step in prior to formal requests for help. Possible approaches could be the integration of novel hardware sensors in complex machines to detect potential problems at an early stage or the use of visual computing to identify prospective help-seekers. In this regard, it is worth noting that security personnel already use CCTV to identify potential threats. However, such an approach would present other challenges such as data protection, privacy and ethical issues that would need to be overcome.

With our design case study, we have explored some of the ways in which modern AR technologies might be applicable in remote troubleshooting and expertise-sharing settings. We hope our results will inform further future work within CSCW that is focused upon supporting remote collaboration and, especially, that they will inspire hardware manufactures to deal with the current shortcomings of AR technologies for the support of distributed guidance and help.

### **Acknowledgements**

We would like to thank all the members of the Fab Lab and all the participants in our empirical study and tests. Without your support and involvement, our design case study would not have been possible.

### **Funding**

Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.



## References

- Ackerman, Mark S.; and Thomas W Malone (1990). Answer Garden: A Tool for Growing Organizational Memory. In *COCS '90: Proceedings of the ACM Conference on Office Information Systems, Cambridge, Massachusetts, USA, 25-27 April 1990*. New York, NY, USA: ACM, pp. 31–39.
- Ackerman, Mark S.; Juri Dachtera; Volkmar Pipek; and Volker Wulf (2013). Sharing knowledge and expertise: The CSCW view of knowledge management. *Computer Supported Cooperative Work (CSCW)*, vol. 22, no. 4-6, 2013, pp. 531–573.
- Adcock, Matt; and Chris Gunn (2015). Using Projected Light for Mobile Remote Guidance. *Computer Supported Cooperative Work (CSCW)*, vol. 24, no. 6, 2015, pp. 591–611.
- Adcock, Matt; Stuart Anderson; and Bruce Thomas (2013). RemoteFusion: Real time depth camera fusion for remote collaboration on physical tasks. In *VRCAI '13: Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry, Hong Kong, Hong Kong, 17 -19 November 2013*. New York, NY, USA: ACM, pp. 235–242.
- Adcock, Matt; Dulitha Ranatunga; Ross Smith; and Bruce H. Thomas (2014). Object-based touch manipulation for remote guidance of physical tasks. In *SUI '14: Proceedings of the 2nd ACM Symposium on Spatial User Interaction, Honolulu, Hawaii, USA, 4-5 October 2014*. New York, NY, USA: ACM.
- Azuma, Ronald T (1997). A Survey of Augmented Reality. *Presence Teleoperators & Virtual Environments*, vol. 4, no. August, 1997, pp. 355–385.
- Bauer, Martin; Gerd Kortuem; and Zary Segall (1999). “Where are you pointing at?” A study of remote collaboration in a wearable videoconference system. In *Digest of Papers. Third International Symposium on Wearable Computers, San Francisco, Calif., USA, 18 -19 October 1999*. New York, NY, USA: IEEE, pp. 151–158.
- Billinghurst, Mark (2013). Hands and Speech in Space: Multimodal Interaction with Augmented Reality Interfaces. In *ICMI '13: Proceedings of the 15th ACM on International Conference on Multimodal Interaction, Sydney, Australia, 9-13 December 2013*. New York, NY, USA: ACM, pp. 379–380.
- Bobrow, Daniel G.; and Jack Whalen (2002). Community Knowledge Sharing in Practice: The Eureka Story. *Reflections: The SoL Journal*, vol. 4, no. 2, 2002, pp. 47–59.
- Bowers, John; and David Martin (2000). Machinery in the new factories: interaction and technology in a bank’s telephone call centre. In *CSCW '00: Proceedings of the ACM conference on Computer supported cooperative work, Philadelphia, USA, 2-6 December 2000*. New York, NY, USA: ACM, pp. 49–58.
- Castellani, Stefania; Antonietta Grasso; Jacki O’Neill; and Frederic Roulland (2009). Designing technology as an embedded resource for troubleshooting. *Journal of Computer Supported Cooperative Work*, vol. 18, no. 2–3, 2009, pp. 199–227.
- Cicourel, Aaron V. (1990). The integration of distributed knowledge in collaborative medical diagnosis. In *Intellectual teamwork: social and technological foundations of cooperative work*. Hillsdale, New Jersey, USA: L. Erlbaum Associates Inc.
- Clark, Herbert H.; and Catherine R. Marshall (1981). Definite reference and mutual knowledge. In A. K. Josh, B. L. Webber, and I. A. Sag (Eds.), *Elements of Discourse Understanding*. Cambridge: Cambridge University Press.
- Crabtree, Andy; Jacki O’Neill; Peter Tolmie; Stefania Castellani; Tommaso Colombino; and Antonietta Grasso (2006). The practical indispensability of articulation work to immediate and remote help-giving. In *CSCW '06: Proceedings of the 20th anniversary conference on computer supported cooperative work, Banff, Alberta, Canada, 4-8 November 2006*. New York, NY, USA: ACM, pp. 219–228.
- Dengel, Andreas (2016). Digital Co-Creation and Augmented Learning. In *KMO '16: Proceedings of the The 11th International Knowledge Management in Organizations Conference on The*

- Changing Face of Knowledge Management Impacting Society, Hagen, Germany, 25-28 July 2016*. New York, NY, USA: ACM, pp. 1–8.
- Elliott, Stephen; and Joan Littlefield (1995). *Educational Psychology: Effective Teaching, Effective Learning*. Brown & Benchmark.
- Ens, Barrett; Joel Lanir; Anthony Tang; Scott Bateman; Gun Lee; Thammathip Piumsomboon; and Mark Billingham (2019). Revisiting collaboration through mixed reality: the evolution of groupware. *International Journal of Human-Computer Studies*, vol. 131, no. Nov 2019, 2019, pp. 81–98.
- Fakourfar, Omid; Kevin Ta; Richard Tang; Scott Bateman; and Anthony Tang (2016). Stabilized Annotations for Mobile Remote Assistance. In *CHI '16: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, California, USA, 7-12 May 2016*. New York, NY, USA: ACM, pp. 1548–1560.
- Fussell, S. R.; R. E. Kraut; and J. Siegel (2000). Coordination of communication: Effects of shared visual context on collaborative work. In *CSCW '00: Proceedings of the ACM conference on Computer supported cooperative work, Philadelphia, USA, 2-6 December 2000*. New York, NY, USA: ACM, pp. 21–30.
- Fussell, Susan R.; Leslie D. Setlock; Jie Yang; Jiazi Ou; Elizabeth Mauer; and Adam D.I. Kramer (2004). Gestures over video streams to support remote collaboration on physical tasks. *Human-Computer Interaction*, vol. 19, no. 3, 2004, pp. 273-309.
- Gao, Lei; Huidong Bai; Gun Lee; and Mark Billingham (2016). An oriented point-cloud view for MR remote collaboration. In *SA '16: Proceedings of ACM SIGGRAPH ASIA Mobile Graphics and Interactive Applications, Macao, Macao, 6-8 December 2016*. New York, NY, USA: ACM, pp. 1–4.
- Gauglitz, Steffen; Cha Lee; Matthew Turk; and Tobias Höllerer (2012). Integrating the Physical Environment into Mobile Remote Collaboration. In *MobileHCI '12: Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services, San Francisco, Calif., USA, 21-24 September 2012*. New York, NY, USA: ACM, pp. 241–250.
- Gauglitz, Steffen; Benjamin Nuernberger; Matthew Turk; and Tobias Höllerer (2014). World-stabilized annotations and virtual scene navigation for remote collaboration. In *UIST '14: Proceedings of the 27th annual ACM symposium on User interface software and technology, Honolulu, Hawaii, USA, 5-8 October 2014*. New York, NY, USA: ACM, pp. 449–459.
- Gavish, Nirit; Teresa Gutierrez; Sabine Webel; Jorge Rodriguez; and Franco Tecchia (2011). Design Guidelines for the Development of Virtual Reality and Augmented Reality Training Systems for Maintenance and Assembly Tasks. In *BIO Web of Conferences, The International Conference SKILLS 2011, Montpellier, France, 15-16 December 2011*. EDP Sciences.
- Gurevich, Pavel; Joel Lanir; Benjamin Cohen; and Ran Stone (2012). TeleAdvisor: A Versatile Augmented Reality Tool for Remote Assistance. In *CHI '12: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, Texas, USA, 5-10 May 2012*. New York, NY, USA: ACM, pp. 619–622.
- Henderson, Steven J.; and Steven K. Feiner (2011). Augmented reality in the psychomotor phase of a procedural task. In *ISMAR '11: Proceedings of the 10th IEEE International Symposium on Mixed and Augmented Reality, Basel, Switzerland, 26-29 October 2011*. New York, NY, USA: IEEE, pp. 191–200.
- Henrysson, Anders; Mark Billingham; and Mark Ollila (2005). Face to Face Collaborative AR on Mobile Phones. In *ISMAR '05: Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality, Vienna, Austria, 5-8 October 2005*. Washington, DC, USA: IEEE Computer Society, pp. 80–89.
- Hoffmann, Sven; Aparecido Fabiano Pinatti de Carvalho; Darwin Abele; Marcus Schweitzer; Peter Tolmie; and Volker Wulf (2019). Cyber-Physical Systems for Knowledge and Expertise Sharing in Manufacturing Contexts: Towards a Model Enabling Design. *Computer Supported Cooperative*

- Work (CSCW): *The Journal of Collaborative Computing and Work Practices*, vol. 28, no. 3–4, May 2019, pp. 469–509.
- Huang, Weidong; and Leila Alem (2011). Supporting hand gestures in mobile remote collaboration: A usability evaluation. In *BCS-HCI '11: Proceedings of 25th BCS Conference on Human Computer Interaction, Newcastle-upon-Tyne, UK, 4-8 July 2011*. New York, NY, USA: ACM, pp. 211–216.
- Huysman, Marleen; and Volker Wulf (2005). IT to support knowledge sharing in communities, towards a social capital analysis. *Journal of Information Technology*, vol. 21, 2005, pp. 40–51.
- Jasche, Florian; and Thomas Ludwig (2020). PrintARface: Supporting the Exploration of Cyber-Physical Systems through Augmented Reality. In *NordiCHI '20: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society, Tallinn, Estonia, 25-30 October 2020*. New York, NY, USA: ACM, pp. 1–12. <https://doi.org/10.1145/3419249.3420162>
- Kim, Seungwon; Mark Billinghurst; Chilwoo Lee; and Gun Lee (2018). Using freeze frame and visual notifications in an annotation drawing interface for remote collaboration. *KSII Transactions on Internet and Information Systems*, vol. 12, no. 12, 2018, pp. 6034–6056.
- Kim, Seungwon; Gun Lee; Weidong Huang; Hayun Kim; Woontack Woo; and Mark Billinghurst (2019). Evaluating the Combination of Visual Communication Cues for HMD-based Mixed Reality Remote Collaboration. In *CHI '19: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4-9 May 2019*. New York, NY, USA: ACM, pp. 1–13.
- Kirk, David S.; and Danaë Stanton Fraser (2006). Comparing remote gesture technologies for supporting collaborative physical tasks. In *CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montréal, Canada, 22-27 April 2006*. New York, NY, USA: ACM, pp. 1191–1200.
- Kristoffersen, S.; J. O'Brien; M. Rouncefield; and J. Hughes (1996). When mavis met iris: ending the love affair with organizational memory. In *IRIS '19: Proceedings of the 19th Information Systems Research Seminar in Scandinavia, Lökeberg, Sweden, 10-13 August 1996*. Göteborg: Department of Informatics, Göteborg University, pp. 29–53.
- Kuutti, Kari; and Liam J. Bannon (2014). The turn to practice in HCI: towards a research agenda. In *CHI '14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Toronto, Ontario, Canada, 26 April - 1 May 2014*. New York, NY, USA: ACM, pp. 3543–3552.
- Ley, B.; T. Ludwig; V. Pipek; D. Randall; C. Reuter; and T. Wiedenhofer (2014). Information and Expertise Sharing in Inter-Organizational Crisis Management. *Computer Supported Cooperative Work: The Journal of Collaborative Computing (CSCW)*, vol. 23, no. 4–6, 2014, pp. 347–387.
- Ludwig, Thomas; Oliver Stickel; Alexander Boden; and Volkmar Pipek (2014). Towards Sociable Technologies: An Empirical Study on Designing Appropriation Infrastructures for 3D Printing. In *DIS '14: Proceedings of the 2014 conference on Designing Interactive Systems, Vancouver, Canada, 21-25 June 2014*. New York, NY, USA: ACM, pp. 835–844.
- Ludwig, Thomas; Alexander Boden; and Volkmar Pipek (2017). 3D Printers as Sociable Technologies: Taking Appropriation Infrastructures to the Internet of Things. *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 24, no. 2, 2017.
- Ludwig, Thomas; Michael Döll; and Christoph Kotthaus (2019). “The Printer is Telling Me About Itself”: Supporting the Appropriation of Hardware by Using Projection Mapping. In *DIS '19: Proceedings of the 2019 on Designing Interactive Systems Conference, San Diego, USA, 23-28 June 2019*. New York, NY, USA: ACM, pp. 331–344.
- Luff, Paul; Christian Heath; Hideaki Kuzuoka; Jon Hindmarsh; Keiichi Yamazaki; and Shinya Oyama (2003). Fractured Ecologies: Creating Environments for Collaboration. *Human-Computer Interaction*, vol. 18, no. 1–2, June 2003, pp. 51–84.
- Lukosch, Stephan; Mark Billinghurst; Leila Alem; and Kiyoshi Kiyokawa (2015). Collaboration in Augmented Reality. *Computer Supported Cooperative Work: The Journal of Collaborative Computing (CSCW)*, vol. 24, 2015, pp. 515–525.

- McDonald, David W.; and Mark S. Ackerman (1998). Just talk to me: a field study of expertise location. In *CSCW '98: Proceedings of the ACM conference on Computer supported cooperative work, Seattle, WA, USA, 14-18 November 1998*. New York, NY, USA: ACM, pp. 315–324.
- Milgram, Paul; and Fumio Kishino (1994). A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information Systems*, vol. 77, no. 12, 1994, pp. 1321–1329.
- Nielsen, Jakob (1993). *Usability Engineering*. Amsterdam etc.: Morgan Kaufmann.
- Normark, Maria; and David Dave Randall (2005). Local Expertise at an Emergency Call Centre. In H. Gellersen, K. Schmidt, M. Beaudouin-Lafon, and W. Mackay (Eds.), *ECSCW '09: Proceedings of the European Conference on Computer Supported Cooperative Work, Paris, France, 18-22 September 2005*. Dordrecht, Netherlands: Springer Netherlands, pp. 347–366.
- O'Neill, Jacki; Stefania Castellani; Antonietta Grasso; Frederic Roulland; and Peter Tolmie (2005). Representations Can be Good Enough. In H. Gellersen, K. Schmidt, M. Beaudouin-Lafon, and W. Mackay (Eds.), *ECSCW '09: Proceedings of the European Conference on Computer Supported Cooperative Work, Paris, France, 18-22 September 2005*. Dordrecht, Netherlands: Springer Netherlands, pp. 267–286.
- Oda, Ohan; Mengu Sukan; Steven Feiner; and Barbara Tversky (2013). Poster: 3D referencing for remote task assistance in augmented reality. In *3DUI '13: IEEE Symposium on 3D User Interfaces, Orlando, FL, USA, 16-17 March 2013*. New York, NY, USA: IEEE, pp. 179–180.
- Orlikowski, Wanda J; and J Debra Hofman (1996). An Improvisational Model of Change Management: The Case of Groupware Technologies. *Sloan Management Review*, vol. 38, 1996, pp. 11–21.
- Orr, Julian E. (1996). *Talking about machines: An ethnography of a modern job*. Ithaca, NY: Cornell University Press.
- Piirainen, Kalle A.; Gwendolyn L. Kolfshoten; and Stephan Lukosch (2012). The joint struggle of complex engineering: A study of the challenges of collaborative design. *International Journal of Information Technology and Decision Making*, vol. 11, no. 6, 2012, pp. 1–39.
- Pipek, Volkmar; Volker Wulf; and Aditya Johri (2012). Bridging Artifacts and Actors: Expertise Sharing in Organizational Ecosystems. *Computer Supported Cooperative Work: The Journal of Collaborative Computing (CSCW)*, vol. 21, no. 2–3, 2012, pp. 261–282.
- Poelman, Ronald; Oytun Akman; Stephan Lukosch; and Pieter Jonker (2012). As if Being There: Mediated Reality for Crime Scene Investigation. In *CSCW '12: Proceedings of the ACM Conference on Computer-Supported Cooperative Work, Seattle, WA, 11 - 15 February 2012*. New York, NY, USA: ACM, pp. 1267–1276.
- Ryle, Gilbert (1945). Knowing How and Knowing That. *Proceedings of the Aristotelian Society*, vol. 46, 1945, pp. 1–16.
- Tait, Matthew; and Mark Billinghurst (2015). The Effect of View Independence in a Collaborative AR System. *Computer Supported Cooperative Work (CSCW)*, vol. 24, no. 6, 2015, pp. 563–589.
- Tecchia, Franco; Leila Alem; and Weidong Huang (2012). 3D helping hands: A gesture based MR system for remote collaboration. In *VRCAI '12: Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry, Singapore, Singapore, 2-4 December 2012*. New York, NY, USA: ACM, pp. 323–328.
- Wang, Peng; Shusheng Zhang; Xiaoliang Bai; Mark Billinghurst; Weiping He; Mengmeng Sun; et al. (2019). 2.5DHANDS: a gesture-based MR remote collaborative platform. *International Journal of Advanced Manufacturing Technology*, vol. 102, 2019, pp. 1339–1353.
- Whalen, Jack; and Erik Vinkhuyzen (2000). Expert systems in (inter) action: diagnosing document machine problems over the telephone. In P. Luff, J. Hindmarsh, and C. Heath (Eds.), *Workplace Studies: Recovering Work Practice and Informing System Design*. Cambridge: Cambridge University Press, pp. 92–140.
- Whalen, Jack; Don H. Zimmerman; and Marilyn R. Whalen (1988). When words fail: a single case analysis. *Social Problems*, vol. 35, no. 4, 1988, pp. 335–362.

- Wobbrock, Jacob O.; and Julie A. Kientz (2016). Research contributions in human-computer interaction. *Interactions*, vol. 23, no. 3, 2016, pp. 38–44.
- Wulf, Volker; Claudia Müller; Volkmar Pipek; David Randall; Markus Rohde; and Gunnar Stevens (2015). Practice-Based Computing: Empirically Grounded Conceptualizations Derived from Design Case Studies. In V. Wulf, K. Schmidt, and D. Randall (Eds.), *Designing Socially Embedded Technologies in the Real-World*. London: Springer London.
- Yamauchi, Yutaka; Jack Whalen; and Daniel G. Bobrow (2003). Information use of service technicians in difficult cases. In *CHI '03: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Ft. Lauderdale, Florida, USA, 5-10 April 2003*. New York, NY, USA: ACM, pp. 81–88.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.