Computers & Graphics 125 (2024) 104095

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Special Section on Advances on SUI 2023

Supporting tailorability in augmented reality based remote assistance in the manufacturing industry: A user study

Troels Rasmussen [a], Kaj Grønbæk [a], Weidong Huang [b,*]

[a] Aarhus University, Denmark
[b] University of Technology Sydney, Australia

ARTICLE INFO

Keywords:
Augmented reality
Remote assistance
Situational awareness
Task space sharing
Guidance location
Hand gestures

ABSTRACT

Research on remote assistance in real-world industries is sparse, as most research is conducted in the laboratory under controlled conditions. Consequently, little is known about how users tailor remote assistance technologies at work. Therefore, we developed an augmented reality-based remote assistance prototype called Remote Assist Kit (RAK). RAK is a component-based system, allowing us to study tailoring activities and the usefulness of tailorable remote assistance technologies. We conducted a user evaluation with employees from the plastic manufacturing industry. The employees configured the RAK to solve real-world problems in three collaborative scenarios: (1) troubleshooting a running injection molding machine, (2) tool maintenance, (3) solving a trigonometry problem. Our results show that the tailorability of RAK was perceived as useful, and users were able to successfully tailor RAK to the distinct properties of the scenarios. Specific findings and their implications for the design of tailorable remote assistance technologies are presented. Among other findings, requirements specific to remote assistance in the manufacturing industry were discussed, such as the importance of sharing machine sounds between the local operator and the remote helper.

1. Introduction

Employees from the manufacturing and field service industries are lead users of remote assistance technology [1]. This is because there are a number of remote assistance working scenarios in these industries. Employees and their organizations benefit greatly from the implementation and use of such technologies. Remote assistance typically involves a remote helper guiding a local worker in the performance of physical tasks [2], such as a remote technician guiding a local worker to fix a broken machine. Technologies that support remote assistance promise to save companies the travel expenses of sending their employees on location and eliminate the need for employees to spend time away from their families. Yet, little is known about how real-world users from industry adapt and tailor remote assistance technology during work, since most research on the topic is conducted in the laboratory under controlled settings.

The research reported in this paper concerns the topic of tailorable remote assistance. It is motivated by the lack of research taking place in an industrial context and informed by a previous study about remote assistance practices in the manufacturing industry [3]. This user study indicates that users will benefit from tailorable remote assistance due to their heterogeneous needs for interface mobility and ways of capturing the task space. More specifically, we conducted a user study investigating users' (remote helpers' and local workers') tailoring work with a component-based remote assistance solution, Remote Assist Kit (RAK). Users can compose RAK configurations from function-oriented components. Helpers can tailor interface mobility (PC or tablet/smartphone) and guidance format (drawings or hand gestures) dimensions of remote guidance, while workers can tailor guidance location (on display or projected into the task space) and task capturing (tablet/smartphone camera or external camera(s)) dimensions. In the user study, we seek to understand the following three aspects:

1. whether users in a manufacturing industry context find tailorable remote assistance useful,
2. whether they will tailor a remote assistance solution, such as RAK, to the distinct properties of problem scenarios, and
3. how they tailor.

Our user study is an experimental simulation [4], with three realistic scenarios using RAK as the vehicle for experimentation. The study was conducted at a craftsmen school for toolmakers in the plastic manufacturing industry, and the participants were students of the school. Unlike most research on remote assistance, our study took

* Corresponding author.
E-mail addresses: troelsrasmussen@outlook.dk (T. Rasmussen), kgronbak@cs.au.dk (K. Grønbæk), weidong.huang@uts.edu.au (W. Huang).

https://doi.org/10.1016/j.cag.2024.104095
Received 2 February 2024; Received in revised form 17 June 2024; Accepted 24 September 2024
Available online 16 October 2024

place in the field, sacrificing precision and control for realism. Instead of systematically controlling an independent variable, which is often the functionality offered by the prototype (e.g. guidance format), we provided our participants with all of RAK's functionalities and allowed them to select and use the features they deemed useful in each scenario, similar to the approach in [5–7].

The results of our user study are divided into two categories: tailoring of remote assistance and insights from the uncontrolled manufacturing industry context where data was gathered. We found that participants perceived the tailorability of RAK as useful and frequently used the same configurations in similar scenarios. We identified patterns in their tailoring work, primarily concerning guidance format and task space capturing. Additionally, we observed unexpected tailoring work beyond the designed capabilities of RAK, and discovered that awareness of the collaborating partner's functional composition of RAK was important. Moreover, we found that sharing machine sounds between a worker and a remote helper is crucial for effective troubleshooting of running machines in production, highlighting a specific requirement for remote assistance in the manufacturing industry.

The rest of the paper is structured as follows: First, we discuss related work on remote assistance, tailorable systems, and their combination. We then describe the design and implementation of RAK, its components, and the type of tailorability it offers. Next, we outline the methodology and procedure for our experimental simulation. Finally, we present and discuss our results, the limitations of our study, and future work.

2. Related work

We intended to implement RAK, a software prototype that is tailorable by end-users. First, we describe related work on remote assistance, including research on augmented reality, guidance formats, guidance locations, and task space sharing, as this research inspired the design of RAK's components. Next, we discuss related work on end-user tailorability in tailorable computer-supported cooperative work (CSCW) systems and previous studies on remote assistance in industry. Finally, we conclude this section with our hypotheses.

2.1. Guidance techniques and task space sharing for remote assistance

A remote guidance technique allows a remote helper to provide instructions to a worker in their shared visual view or space. Traditional remote guidance methods include audio-only or video conference setups [8]. Kraut et al. [9] conducted a study in which different audio and video configurations were tested and compared. Particularly, in their video conditions, both the helper and worker sides have a camera to capture the local scene so that the local view can be shared with each other for collaboration. A similar study was also conducted by Fussell [10] showing the value of a shared view of the work environment, or shared visual space, for remote guidance on physical tasks.

Later with the advancement of emerging technologies including augmented reality and mobile devices, the instructions provided by the remote helper were often implemented through augmented reality techniques [2]. Augmented reality is a technology that integrates computer-generated digital information such as texts, visuals and sounds with the physical environment in real time providing interactive experience for end users. For remote assistance, the helping information is often displayed in the shared visual space and used to augment with the real-world workspace of the worker. For example, the DOVE system of Fussell et al. [11] allows the helper to draw sketches on the video of the worker's workspace. The combined digital sketches and the real-world scene information is then displayed to the worker. Huang et al. [12] extracted unmediated hand gestures from the helper side video and combined them with the video of the worker's workspace in their HandsInAir mobility system. Further, Huang et al. [13] combined both hand gestures and sketches with the scene of the worker's workspace in HandsInTouch for more advanced collaboration.

Guidance techniques have a format (e.g. pointer, drawings, or hand gestures) and location (e.g. non-stabilized drawings on an external monitor separate from the task space, or augmented reality drawings registered to specific locations in the task space). Research on remote guidance techniques often compares the guidance format [11,13–16] or the location of the guidance [14,17,18] with respect to how well the worker–helper pairs collaborate and perform on a given physical task.

For instance, Kirk et al. [14] had worker–helper pairs go through an assembly task, comparing freehand sketching to unmediated hand gestures. They found a significant difference in model completion, concluding that there were performance gains from using an unmediated representations of hands. In the same study, Kirk et al. also compared the output locations of the helper's gestures: external monitor vs. projecting gestures on the table surface of the task space. They did not find a significant difference in performance between the locations. Benefits of combining drawings and hand gestures in a shared view have also been demonstrated [13]. Some researchers have also investigated how drawings [11,18] or hand gestures [19] are used by categorizing them in terms of their communicative function.

The setup of cameras in a task space and the camera technology determine how the worker's task space is captured and shared with the helper. An important concept to consider in the design of task space sharing is view independency, i.e., the helper's ability to access a task space independently of the worker's movements. Several research projects have shown that view independency of the helper is beneficial to collaboration [8,20–25]. This is because the remote helper can more easily understand a problem at their own pace and provide guidance when access to the task space is independent of the worker's movements.

Based on the related work discussed above, we have the following two design requirements for RAK:

1. RAK supports components for changing the guidance format (drawings or hands) and guidance location (on external display or projected into the task space).
2. Through the task space capturing components, RAK supports both view dependence and view independence, with the webcam(s) offering a view independent option and the smartphone/tablet offering both a view dependent and independent option, depending on whether the smartphone/tablet is handheld or mounted in the environment.

Further, prior research [5,7] investigated the purposes and scenarios in which different remote assistance functionalities were used, including when and how collaborators employed various guidance techniques when they could freely choose from all the functionalities in one solution. Our work is related in that our user study also focuses on exploring the full functionality of RAK, rather than systematically comparing guidance techniques or methods of sharing the task space, as is often the case in remote assistance research. With the tailorability of RAK, we aim to investigate whether users find different functionalities useful in different problem scenarios.

2.2. Tailorable CSCW systems

End-user tailoring of software applications has been explored extensively. Component-based tailoring involves end-users composing applications from reusable software components [26]. Mørch [27] divided end-user tailoring of software applications into three levels: customization, integration, and extension, each with ascending tailoring power. In the context of component-based tailoring, customization involves configuring the infrastructure that contains the components or making a change that is propagated to all components. Integration means composing an application from a set of existing components.

Extension involves adding new programming code to either the infrastructure or developing an entirely new component [26]. Whether end-users can take advantage of the tailoring power offered by the three levels depends on their skills, as tailoring techniques with more tailoring power require proportionally better tailoring skills [28].

The ability of users to tailor CSCW systems has been found to be important due to individual differences among collaborating users, such as different preferences, experiences, and skills, and because different collaborative tasks pose varying requirements for functionality [26,29,30]. Important design considerations for tailorable CSCW systems include the ability to tailor the system during run-time, to avoid having to shut down the system for all users because one user is tailoring parts of it, tailoring rights, and scoping the tailoring activities to individual and group scopes [30]. Run-time tailorability of CSCW systems poses a challenge if a tailoring step by one user affects other users (depending on the scope of the activity). To remedy this problem, Herrmann et al. [31] proposed the design principle of negotiability. Specifically, they implemented a feature in their tailorable CSCW system that allowed users to reject, accept, or modify a tailoring step made by another user.

RAK can be perceived as a tailorable CSCW system, where tailoring (integration) work can happen during run-time. The tailorability of RAK is an example of pure integration; the remote assistance application is composed of a set of components, and no information is required to specify how the components are connected. For instance, a helper can choose to replace one guidance format with another during a remote assistance session. The helper's choice of guidance format module naturally affects the worker, who is the receiver of the guidance. Similarly, the worker's choice of task space sharing component will affect the helper's visual access to the task space and may lead the helper to reconsider which guidance format or mobility components are best suited for the current task space sharing component. Therefore, we have implemented negotiability in the design of RAK and study how RAK supports negotiation during the composition of configurations.

To our knowledge, only one research example of tailorable remote assistance exists: the work by Speicher et al. [7]. They evaluated a component-based tailorable remote assistance system which utilized a non-tailorable rigid core functionality and function-oriented components. These components could be selected and deselected via a graphical user interface. The system provided support for multiple collaborators, 360-degree cameras, projected guidance into the task space, drawing on the 360-degree video, session persistence, and video rewind features. Similarly, RAK also supports projected guidance. However, RAK differs by offering various guidance formats such as drawing or hand gestures, and it supports different methods for capturing the task space, including tablet/smartphone cameras or one or more external cameras.

To summarize, we have the following design requirements for RAK as a tailorable CSCW system:

1. RAK implements negotiability and supports negotiation during the composition of configurations.
2. RAK supports projected guidance and provides different guidance formats and various camera capturing methods.

2.3. Industrial remote assistance

The majority of research on remote assistance involves laboratory studies, where a remote helper must guide a worker through a stylized task, such as the assembly of abstract LEGO models or tangram puzzles.

Comparatively, studies on remote assistance within industry – whether they are short field experiments, experimental simulations, or longitudinal field studies – are surprisingly sparse. Several research papers have documented case studies where remote assistance solutions were designed and sometimes evaluated specifically for industries such as manufacturing [6], office environments [32], healthcare [5,33], mining [34], or security [35]. For instance, Alem et al. [34] developed a remote assistance system tailored for maintenance operators in mines. Their experimental simulation, however, was conducted in an office setting rather than in the actual mines. Nonetheless, they evaluated the system by involving representative end-users in two realistic repair scenarios: repairing a photocopy machine and removing a card from a computer motherboard. Domova et al. [6] evaluated their remote assistance system with real-life technicians and control room operators in a waste water treatment plant. Their participants were required to go through simulated maintenance scenarios that had previously been identified. At the level of realism their user study is similar to ours, because we also had our study participants, toolmakers from the plastic manufacturing industry, go through simulated realistic scenarios, which were previously identified by their peers.

2.4. Hypotheses

In this section, we have reviewed related work on augmented reality based remote assistance systems together with relevant guidance techniques and tailorable CSCW systems. We also derived some design requirements to explore tailorability of RAK with the aim of addressing the following hypotheses:

1. H1: Users will tailor remote assistance to the requirements of a problem. Therefore, with respect to the preferred configurations there is little variation within each of the remote assistance scenarios, but large variation between scenarios due to their distinct requirements.
2. H2: Users perceive the tailorability of RAK as useful for remote assistance tasks. That is, they perceive a configuration of RAK tailored to a scenario as more useful than a standard video communication solution.

3. Remote Assist Kit (RAK)

Tailorable products offer users the advantage of addressing need-information-intensive subtasks, while allowing manufacturers to handle solution-information-intensive subtasks [36]. This approach stems from the belief that users are best positioned to understand their functional and aesthetic requirements (need information) and should thus be responsible for product design. Conversely, in the realm of software, developers excel in determining technical specifics like programming languages and network protocols (solution information), making them responsible for ensuring the technical platform supports the required end-user tailorability. Empowering users to innovate and customize products to their needs has been shown to enhance user satisfaction, as observed in domains such as security software [37]. We argue that this concept applies to remote assistance solutions as well, supported by our review of related work in Section 2 and a previous interview study in the manufacturing industry, where users expressed diverse needs related to interface mobility and task space capture [3].

Remote Assist Kit (RAK) is a customizable remote assistance solution that allows users – comprising a remote helper and a local worker – to configure interface mobility, guidance format, guidance location, and task space capturing according to their specific needs and preferences. RAK consists of three JavaScript-based web apps: the Helper App, Worker App, and Supplementary Worker App, designed to run seamlessly on both PCs and tablet/smartphone devices. Communication between these apps, including two-way video and audio streaming and one-way guidance, is implemented using WebRTC. The Helper App enables the helper to tailor the remote assistance experience in two primary ways. Firstly, they can choose the interface mobility by selecting to provide guidance via video of the worker's task space on either a PC or tablet/smartphone. Secondly, the helper can select the guidance format, opting between drawings or hand gestures. The choice

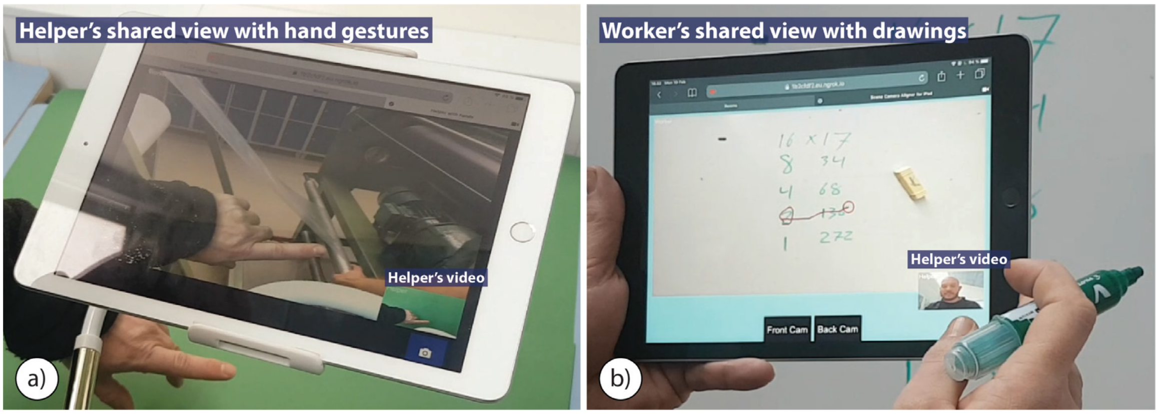

Fig. 1. Interface of the Helper App and Worker App follows the "What You See Is What I See" (WYSIWIS) paradigm. (a) The shared view in the Helper App displays the helper's hand gestures overlaid on the video of the worker's task space. This view is shared with the Worker App (not pictured here). (b) The shared view in the Worker App shows the helper's drawings superimposed on the video of the task space. This view is shared with the Helper App (not pictured here).

of guidance format determines the required input devices: a mouse for drawing on a PC, touch input for drawing on a tablet/smartphone, and an external webcam for capturing hand gestures on a PC, or the integrated back-facing camera on a tablet/smartphone.

The Worker App allows the local worker to customize task space capturing. This can be achieved using one or more webcams connected to a PC or the built-in camera of a tablet/smartphone. Multiple webcams provide flexibility in covering different areas or perspectives of the task space. The live video feed from these webcams and audio from the PC microphone are shared with the helper, who can select which camera view to observe from a list. This setup offers the helper view independence, enabling them to monitor various aspects of the task space independently of the worker's location. Alternatively, using a tablet/smartphone camera for task space capturing allows the worker untethered mobility.

Both the Helper App and Worker App interfaces follow the "What You See Is What I See" (WYSIWIS) paradigm. This means both the helper and worker see a shared view of the worker's task space, overlaid with the helper's guidance. Specifically, the helper's drawings are superimposed on top of the live video of the task space. For hand gestures, the helper's hands are extracted from their video feed using a "green screen" (uniformly colored background) and alpha blending technique, similar to the method described in Huang et al. [38] (see Fig. 1).

The Supplementary Worker App allows a worker to optionally customize the guidance location by selecting to view a helper's guidance on an additional tablet/smartphone display or projected into the task space using a projector connected to a PC. On a tablet/smartphone, the interface follows the WYSIWIS paradigm, displaying a shared view of the worker's task space with the helper's guidance overlaid. However, when guidance is projected into the task space, the helper and worker view the task space and guidance from different perspectives, thus the projection interface does not adhere to the WYSIWIS paradigm. The projected guidance provides a spatial augmented reality experience, akin to techniques described in [23]. It is important to note that in the Worker App, the helper's guidance always appears in a shared view of the task space, regardless of whether the Supplementary Worker App is in use or not.

Importantly, these tailoring dimensions – interface mobility, guidance format, task space capturing, and additional guidance location – are independent of each other. This means a helper's choices regarding interface mobility and guidance format do not affect a worker's choices regarding task space capturing and additional guidance location.

Users of RAK have access to a variety of mounts such as gooseneck mounts, gorillapods, tripods, and magnetic mounts. These mounts offer customization opportunities within the interface mobility dimension for the helper: the tablet or external camera of the PC must either be handheld or securely mounted in the workspace. This choice affects the helper's ability to use one or two hands for guidance. Similarly, the mounts provide tailoring options within the task space capturing dimension for the worker: the tablet/smartphone or webcam(s) connected to a PC can be handheld or mounted in the workspace.

When using a handheld camera, the worker can only use one hand to manipulate objects in the task space, resulting in potentially unstable views for the helper and offering multiple perspectives on task space objects. In contrast, a mounted camera allows the worker to use both hands for manipulation, provides a stable view for the helper, and limits the perspectives available on task space objects. If workers choose to mount one or more external webcams, they must consider various surface mounting options in the environment to ensure effective coverage of task space areas. In both handheld and mounted scenarios, camera positioning – considered a specialized form of tailoring – is essential to ensure the helper obtains an appropriate view of the task space.

The combinatorial landscape of components in RAK – encompassing devices, guidance formats, and mounts – is illustrated in Fig. 2, resulting in a total of 324 unique RAK configurations. In the interface mobility dimension, a helper selects one device and mount option, alongside choosing between drawings or hand gestures for guidance format. Similarly, in the task space capturing dimension, a worker selects one device and mount option. Optionally, in the additional guidance location dimension, the selection of a device is available. Importantly, the selection of components in one dimension is independent of selections in another, and these choices are concatenated to uniquely describe each configuration. During our experimental simulation, we examined users' tailoring processes, including their configuration preferences, camera setups, and choice of guidance formats (see Section 5).

4. The user study

We conducted a user study with RAK to understand whether users will tailor remote assistance solutions to the distinct properties of problem scenarios, and how they tailor remote assistance. The user study was an experimental simulation [4], meaning that RAK was used by professional toolmakers from the plastic manufacturing industry in simulated realistic problem scenarios.

4.1. Scenarios

To get the students to explore the tailorability of RAK, we had them go through three realistic remote assistance scenarios, which we hypothesized would pose varying requirements to remote assistance due

	Helper			Worker		
No.	Interface mobility		Guidance format	Taskspace capturing		Additional guidance location (optional)
	Device	Mount		Device	Mount	Device
1	Tablet/smartphone	Tripod	Drawings	Tablet/smartphone camera	Tripod	Tablet/smartphone display
2	PC	Gooseneck	Hand gestures	Single webcam (PC)	Gooseneck	Projection (PC + projector)
3		Gorillapod	Verbal only	Multiple webcams (PC)	Gorillapod	
4		Handheld			Handheld	
	Example configuration "11223"					
	Tablet/smartphone	Tripod	Hand gestures	Single webcam (PC)	Gorillapod	

Fig. 2. Hardware (devices and mounts) and software (guidance formats) component-based tailorability of RAK. Table shows the components in RAK's tailoring dimensions. A *RAK configuration* consists of a component from each column (unless it is optional). The numbers of the components are concatenated and used to uniquely describe the configurations as demonstrated with the example configuration "11223".

Fig. 3. The three scenarios picked in collaboration with a teacher from the craftsmen school: (a) trigonometry exercise, (b) troubleshooting of running injection molding machine, (c) tool maintenance.

to their distinct properties and thus bring the different configurations of RAK in play. In collaboration with the main responsible teacher, who is a specialist in plastic manufacturing, we identified these three scenarios: (1) Trigonometry exercise, (2) Troubleshooting of running injection molding machine, (3) Tool maintenance (see Fig. 3). The teacher crafted the problem–solution space for each scenario and was therefore responsible for preparing locations, tools and machines. To make it easier for the teacher to design scenarios, we made sure that he had experience with RAK by conducting a workshop with him and some other teachers, where they got to play around with the tailorability of RAK on the school's machines.

4.1.1. Scenario 1 - Trigonometry exercise

A worker must solve two trigonometry exercises in collaboration with a remote helper. For the first exercise the worker is provided an illustration of a right-angled triangle, where one angle and the opposite side is given. The worker must find the remaining angles and sides of the triangle in collaboration with the helper. For the second exercise, the worker is provided a written description of a triangle, where two sides and one angle is given. It cannot be assumed that the triangle is right-angled. Again, the worker must find the remaining angles and sides of the triangle in collaboration with the helper. For both exercises only the helper has access to a sheet with trigonometry formulas. We henceforth call the scenario by the name "trigo"-scenario.

4.1.2. Scenario 2 - Troubleshooting of running injection molding machine

A running injection molding machine produces plastic pieces that contain imperfections. A worker therefore needs to troubleshoot the machine in collaboration with a remote helper. The problem is solved when the pieces look as expected without any imperfections. The solution, not provided to the worker or helper, is to adjust input parameters on the human–machine interface. We henceforth call the scenario by the name "injection"-scenario.

4.1.3. Scenario 3 - Tool maintenance at the workbench

A mold (used inside an injection molding machine) at the workbench needs maintenance. A worker needs to disassemble the mold, lubricate parts inside of the mold, and reassemble the mold. A helper guides the worker through the disassembly and reassembly process. The task is solved, when the mold is correctly reassembled. We henceforth call the scenario by the name "tool"-scenario.

4.2. Participants

We recruited twelve toolmaker students aged between 21 and 43 (average was 33 years) from a craftsmen school. They were all from the same top-up education, the plastic manufacturing specialist education, and the same class. These twelve students made up the entire class. All of the students worked as professionals in the plastic manufacturing industry and had an average of 11 years of work experience. The students were grouped in six pairs for the evaluation. The worker–helper pairs were created by the main responsible teacher, who also helped setting up the scenarios, and were based on the groups they would normally use for class work. Thus, the pairs were used to working together. The participants would rate their prior experience with a scenario after its completion. A pairwise comparison of the worker–helper pairs' experience with the scenarios (see Fig. 4) reveals that workers and helpers rated themselves as having approximately the same experience in the "injection" and "tool" scenarios. Half of the workers had little perceived experience with the "trigo"-scenario compared to their helper (G2, G3, G4), while both parties in G1 had little experience. Additionally, the helper-participants were given a sheet with trigonometry formulas (but not the exact solution to the trigonometry problems) to ensure an asymmetrical knowledge-relationship

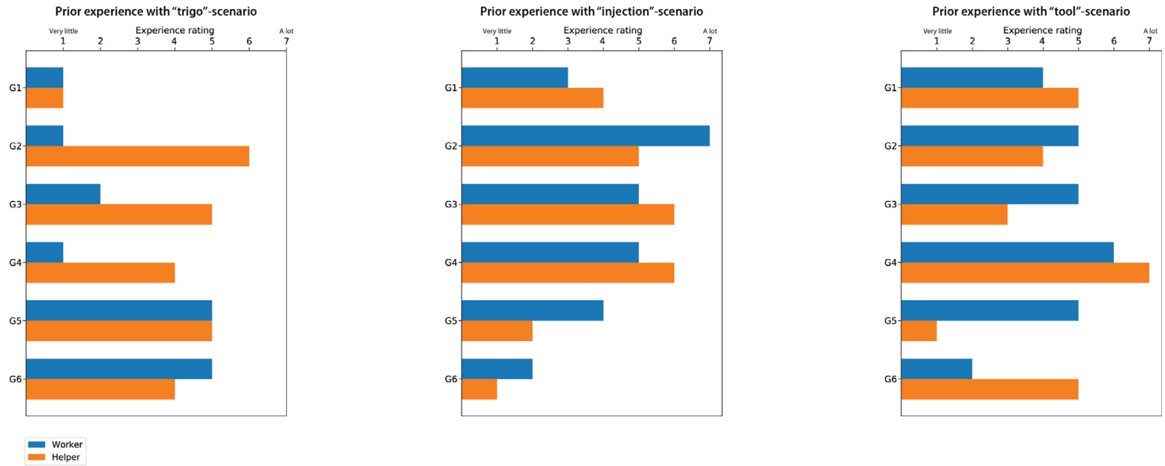

Fig. 4. Workers' and helper's perceived prior experience with each scenario.

4.3. Setup and equipment

The experimental simulation with RAK required various equipment, including smartphones, tablets, web cameras, a laptop, a projector, and various smartphone/tablet/camera/projector mounts to support the different configurations. The worker and remote helper were each provided a large bag with equipment. The worker's bag contained one tablet, one smartphone, one laptop, three webcams, one small portable projector, two gorillapods for the webcams or projector, one gorillapod for the tablet, one tripod for the tablet, one gorillapod for the smartphone, and one flexible gooseneck mount for the tablet/smartphone. The helper's bag contained one tablet, one laptop, one webcam, one tripod for the tablet, and one gorillapod for the webcam.

Besides the equipment required to use RAK, the worker and helper required pen and paper for writing calculations in the "trigo"-scenario, the "injection"-scenario required an injection molding machine, and the "tool"-scenario required a mold and various tools (hammer, screwdriver, Allen key and wrench) for disassembling and reassembling it.

The setup required the remote helper to sit or stand (depending on choice of configuration) in an office space separate from the worker. Three working areas were created in the office space: green pieces of carton were placed on the surface of a meeting table and on a TV and acted as green screens, in case a helper choose hand gestures as guidance format. Additionally, a whiteboard next to the TV was used as is. The setup for the worker depended on the scenario. In the "injection"-scenario he would do work standing at an injection molding machine. In the "tool"-scenario, he would sit at a workbench in the school's workshop space. In the "trigo"-scenario he would sit at a table or stand at a whiteboard in an office space that was similar to the space occupied by the remote helper. It was part of the participants' tailoring work to decide on a setup for the task space capturing devices (external webcam(s) / integrated tablet camera) on the worker's site and the guidance input devices (touch display/external webcam/integrated tablet camera) on the helper's site by using the mounting equipment before and during remote assistance.

4.4. Procedure

Students arrived to the meeting point, an office space at the craftsmen, in pairs. First, they filled out a short pre-study questionnaire about their age, work experience in the plastic manufacturing industry, and previous experience with remote assistance in the role of either the local worker, remote helper or both.

Then, RAK and the three experimental scenarios were explained to the pairs using a slideshow presentation. The tailorable dimensions of RAK (guidance input, guidance format, task space capturing, and guidance location) and its different configurations were illustrated with a system diagram. Special attention was paid to explaining that they had two tasks: one was of course to solve the problem in each scenario, but an equally important task was to find the most useful configuration for a given scenario, rather than solving the problem as quickly as possible. They were told that they were not timed and they would not be judged based on whether they solved a problem correctly or not, as the correct solution was irrelevant to the focus of the study.

Next, some of the configurations of RAK was demonstrated to the pairs by an experimenter, after which the participants could shortly explore the configurations to their liking in an open remote assistance scenario that required assembly of LEGO bricks on a table. In this small training session, the helper and worker were still in the same room, the helper's office space The purpose of the training session was to familiarize students with the functionality of RAK. Afterwards, the pairs assigned themselves roles as helper and worker based on prior experience with the first scenario. We explained to them that they had to assign the helper role to the person, who they perceived to have the most experience with the scenario.

Using RAK, they went through the three realistic remote assistance scenarios presented earlier. The order of the scenarios were randomized and counter-balanced. The worker was taken to the workspace involved in the scenario, while the helper stayed in the office space. The pair were no longer able to see or hear each other directly.

For each scenario the pairs were asked to solve a problem and collaboratively explore configurations to identify the most useful ones. To aid the exploration, the content of the equipment bags were laid out on the floor in their respective workspaces. They were allowed to change configurations as many times as they wanted *during* a scenario. This procedure thus differed from that in the evaluation by Speicher et al. [7] in which study participants (university students) had to specify one configuration of a remote assistance prototype and the reasons why they believed the particular configuration was useful, *before* beginning a remote assistance scenario. A maximum of 30 min were provided for each scenario, but the pairs were allowed to opt out before the 30 min mark, if they had explored all the configurations they wanted and solved the problem.

After each scenario, the helper and worker filled out a post-scenario questionnaire. They would then assign themselves new roles as either helper or worker based on prior experience with the next scenario. At the end of all scenarios, they filled out a post-study questionnaire. Finally, we asked participants to think of other scenarios, where they would find RAK useful, and how the would configure RAK to support these scenarios.

4.5. Measurements

Video cameras recorded the pairs. The camera at the worker's location was placed strategically covering both the task space, the worker's manipulations and actions in the space, and the interface of RAK. The camera at the helper's location covered his interactions with the RAK interface. We coded the video recordings to identify common and outlier usage patterns. Specifically, we looked for patterns in the helper's guidance format (drawings, hands, or verbal descriptions), the helper's gestures (pointing, kinetic, or iconic [11]), the worker's camera work (handheld or mounted scene camera(s), close-up view, overview, moving camera in relation to object, moving object in relation to camera), and unexpected occurrences of tailoring work or interface challenges.

Post-scenario questionnaires were provided to the worker and helper after each scenario to measure and rank the perceived usefulness of the RAK configurations. They also wrote additional explanations of why they preferred one or more configurations. A post-study questionnaire was provided at the end to measure the perceived usefulness of the tailorability of RAK.

To systematically analyze the number of occurrences of configurations, we gave each configuration a unique number based on its composition of components. For example, a helper selecting a tablet on a tripod and the option to use hand gestures gives a helper-configuration with the unique number "112" due to the position of the components in the lists in Fig. 2. Now, if the worker chooses to share his task space using a single webcam (connected to a PC) mounted on a gorillapod, then this worker-configuration has the unique number "23". The combined configuration has the unique number "11223".

5. Results

In this section we present the results of our experimental simulation. First we analyze the participants' configuration preferences in Section 5.1. Then we investigate the perceived usefulness of configurations in Section 5.2. The tailoring work of participants unfolded beyond the choice of configurations and included tailoring of guidance formats and camera work, some of which was surprising and unexpected (see Sections 5.3 and 5.4). Finally, we discuss some requirements specific to remote assistance in the manufacturing industry that were identified from observations during the user study and during an interview with the main responsible teacher at the craftsmen school.

5.1. Analysis of configurations

A total of 324 configurations were available. We observed worker–helper pairs used 19 unique configurations (5.9 pct.). However, keeping the low sample size in mind (6 groups), even if each group had explored two unique configurations per scenario, it would only have amounted to 36 unique configurations (6*2*3 = 18), which is still well below 324 configurations. Consequently, a large portion of the configurations were never used, such as configurations with multiple cameras, and some were used very rarely, for instance configurations that included hardware for extra guidance. Groups typically used one configuration per scenario. They were encouraged to explore the configurations of RAK to identify the one(s) that they perceived as most useful for each scenario *during* a scenario, as described in Section 4.4. We did observe one helper explore different mounts, first gooseneck then tripod, during a scenario, however none of the helpers ever changed device during a scenario. Also, contrary to what we expected, the workers did not explore different camera devices and mounts during a scenario, however one worker did add a guidance display device during a scenario. This means that helpers and workers most often settled on a RAK configuration and completed most tailoring work before a scenario began. Therefore, iterative ongoing tailoring work during remote assistance manifested mostly as camera work, where workers would pick up a camera device with a mount and move it or hold it. We speculate that a plausible reason for the limited ongoing tailoring work was that participants tended to highly prioritize and be engaged in the task of problem solving. This was indicated by the disappointment of some of the participants when they were not able to solve a problem within the time frame.

To investigate H1: *With respect to the preferred configurations there is little variation within each of the remote assistance scenarios, but large variation between scenarios due to their distinct properties*, we visualize and look at all configurations that were used two or more times in any of the scenarios (see Fig. 5). Making use of statistical tests to investigate H2 makes no sense due to the small sample size (N = 6 groups). Instead, from the visualization of configuration us (see Fig. 5) we see signs of support of H2, when we look at the workers' and helpers' configurations, but not for the combined configurations. Below, we check for the support of H2 in all combined configurations of worker–helper pairs, helpers' configurations, and workers' configurations.

Combined configurations: Looking at the worker–helper pairs' combined configurations (see Fig. 5, a), there is no clear preference for a configuration in any of the scenarios, so there is no support for H2 for the combined configurations.

Helpers' configurations: The helper's configurations (see Fig. 5, b) were concerned with different guidance formats and mounts, since the interface mobility device, a tablet, remained the same across all scenarios and participants. In the "trigo"-scenario, 3/6 helpers used the tablet with a tripod and drawings on whiteboard ("111"), and 4/6 helpers used the tablet on a gooseneck with drawings ("121"). None of the groups used verbal guidance only. There was a preference for these two configurations ("111" and "121") only in the "trigo"-scenario. In the "injection"-scenario, 3/6 helpers used the tablet on a gooseneck with verbal guidance only ("123"), indicating a slight preference. Most notably, variance within the "injection"-scenario came from the choice of guidance format, because every helper chose to use a tablet on a gooseneck. Also in the "tool"-scenario there was a slight preference for tablet on a gooseneck with verbal guidance only ("123"), which was used by 3/6 helpers. There is no support of H2 for the helpers' configurations. However, the results indicate that the categorically different tasks and the difference in the participants' experience with the "trigo"-scenario on one side and the "injection" and "tool"-scenarios on the other side led to different preferences for guidance formats, with drawings dominating the "trigo"-scenario and verbal guidance dominating the other two scenarios.

Workers' configurations: The worker's configurations (see Fig. 5, c) were concerned with devices and mounts for task space capturing and additional guidance location. When we examine the workers' configurations in the "trigo"-scenario, we see no clear preference for a configuration. 5/6 workers chose tablet on a tripod ("11") in the "injection"-scenario, while no other configurations were used multiple times. Moreover, this configuration ("11") was not used in any of the other scenarios, so it was clearly preferred only for the "injection"-scenario. In the "tool"-scenario, 5/6 workers used a single webcam on a gorillapod ("23") and no other configurations were used multiple times. The same configuration was used two times in the "trigo"-scenario, which indicates that it was preferred only for the "tool"-scenario. In conclusion H2 is only supported for the worker's configurations if we isolate the "injection" and "tool"-scenarios. These two scenarios had distinct properties, namely the size of the task space which was large in "injection"-scenario and small in the "tool"-scenario and therefore resulted in different preferences for how the task space was captured.

With respect to the preferred configurations, there is more variance within each scenario for the combined configurations than for the worker's and helper's configurations. This implies that a worker's choice of configuration does not affect a helper's choice of configuration and vice versa. This is supported by the video observations, where no evidence was found that participants would negotiate the use of a particular configuration.

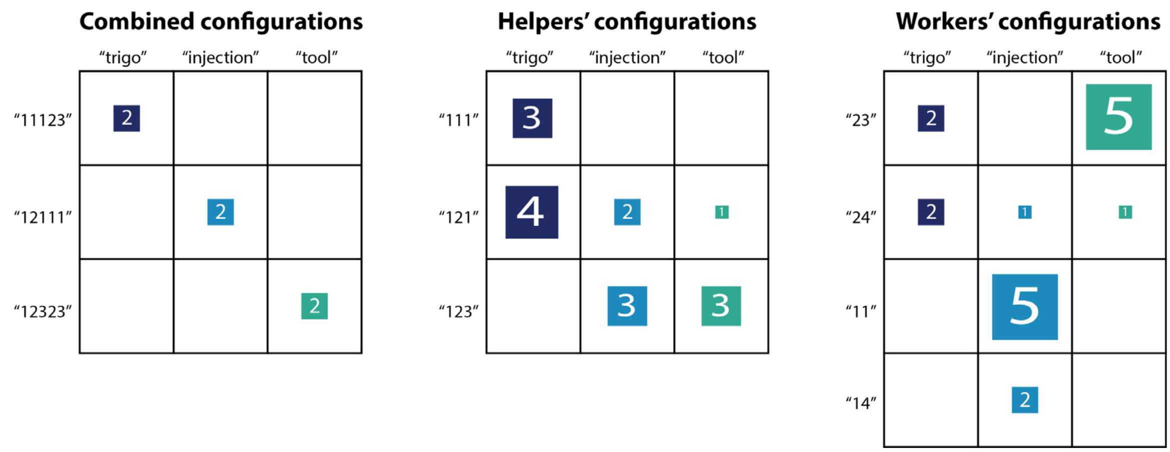

Fig. 5. Configurations that were used two or more times in any of the scenarios. (Left) Worker–helper pairs' combined configurations. (Middle) Helpers' configurations. (Right) Workers' configurations.

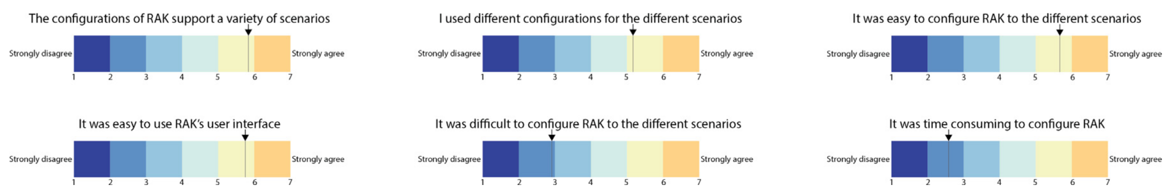

Fig. 6. Participants' average ratings of statements in post-study questionnaire.

5.2. Perceived usefulness of configurations

To investigate H2: *Users perceive the tailorability of RAK as useful*, the participants answered questions regarding the usefulness of the tailorability of RAK in the post-scenario and post-study questionnaires. In the post-scenario questionnaires they rated the statement "RAK had one or more configurations that supported the scenario well", with an average rating of 5.5 in the "trigo"-scenario, 5.8 in the "injection"-scenario, and 5.9 in the "tool"-scenario. Thus, the participants perceived one or more configurations as useful in each scenario. A Mann–Whitney U test showed no difference between the workers and helpers in how well they perceived the configurations to support the scenarios. Additionally, a Friedman test showed no difference in ratings between scenarios. Combined with the previous insight that helpers and workers tailored configurations to the different scenarios, the overall positive rating of the above statement supports H2. In the post-study questionnaire (see Fig. 6) they rated statement 1, "The configurations of RAK support a variety of scenarios", with an average of rating of 5.8. They rated statement 2, "I used different configurations for the different scenarios", with an average rating of 5.2. 6/12 participants rated statement 2 with a 6 or 7 (strongly agree), which indicates that the tailorability was perceived as useful and it thus further lends support to H1. However, we did not only use questionnaire data to measure how well (or ill) suited the chosen configurations were in the scenarios. Video observations and participants' suggestions for improvements show how they coped with and tailored the chosen configurations, which is the subject of the next section.

5.3. Tailoring of guidance formats

We observed that helpers tailored their non-verbal guidance formats (drawings and hand gestures) and verbal communication to the distinct properties of the scenarios.

Drawings were predominantly used in the "trigo"-scenario. Referential iconic drawings [11] were used a lot by the helper to recreate the triangle described in the worker's exercise in her own space or to annotate shared video of the triangle in the worker's space with formulas and calculations. Half of the helpers drew on top of video of the worker's task space, while the other half drew on a whiteboard. Pointing to the sides of the triangle with either the hand or using an arrow was done less frequently, likely because the shared video of the triangle allowed effective communication by just referring to the angles and sides with the names A, B and C.

In comparison to the "trigo"-scenario, in the "injection" and "tool"-scenarios, the pairs already shared a view of the task space and therefore did not have to spend the same effort establishing common ground [39] using drawings. Furthermore, the "tool"-scenario did not warrant much use of the non-verbal guidance formats, because the pairs mostly communicated verbally by using technical terms to refer to the different parts of the mold as they shared the same technical vocabulary. For this reason, we speculate that in the "tool"-scenario verbal communication required less communicative effort of helpers than using the non-verbal guidance formats, indicated by the observation that some helpers would make a conscious attempt at using drawings or hand gestures, but then quickly either decided to revert to verbal communication or simply forgot to use non-verbal guidance. The "injection"-scenario led to different cases: half of the helpers used one of the guidance formats frequently throughout the scenarios, while the remaining helpers used verbal guidance only. Specifically, the guidance formats, both drawing and hands, were used more frequently for pointing in the "injection"-scenario than in the "tool"-scenario. We speculate that this has to do with the objects being referred to in the "injection"-scenario, oftentimes similarly looking buttons on the human–machine interface, being more difficult to describe verbally than the mold parts, which looked distinct and had technical names. If this is the case, this observation is related to the finding by Kraut et al. [40] that a shared view is more useful when users have no precise vocabulary for describing the task space.

In the "trigo"-scenario, a worker–helper pair setup a configuration with a projector, which revealed an interesting interface challenge and requirement to tailorable remote assistance. The helper's drawings were projected onto a whiteboard in the worker's space. However, the helper was not aware of how his guidance looked, as he would write formulas that covered both the whiteboard and the wall behind the whiteboard

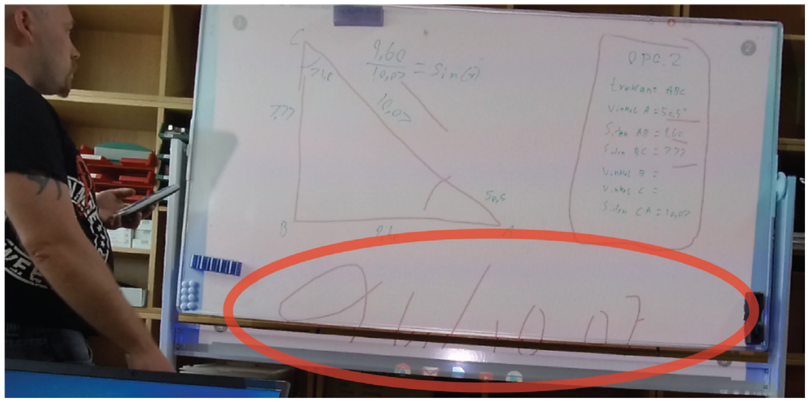

Fig. 7. Helper draws formulas that are partly outside the worker's whiteboard, but is not aware of how unintelligible it looks to the worker.

Fig. 8. Unexpected tailoring work.

(see Fig. 7). To the helper the formulas where shown in 2D on top of video of the task space from a fixed perspective and therefore looked fine. But, unbeknownst to him, the projections on the wall were warped and therefore not readable for the worker. This case illustrates well the importance of having the right mental model of how the projection component visualizes guidance and its limitations, and that the helper needs to be made aware of the worker's selection of additional guidance location component through information in the interface. This problem also exemplifies the importance of worker–helper pairs negotiating how to compose their RAK configuration.

Some participants tailored drawings in an unexpected way in the "trigo"-scenario. Rather than focusing on a shared view of the worker's task space, which RAK supports intentionally, they solved the trigonometry exercises by focusing on a shared view of the *helper's task space*, a whiteboard on which the helper drew the exercises (see Fig. 8, a). Thus, the camera work was inversed, leading the worker to ask the helper to make adjustments to the camera view. Participants took advantage of the hand gesture component's algorithm for segmenting the helper's hands: the algorithm regards everything, which does not fall inside the range of a user-selected background color, as foreground. Since the user-selected background color was green at the time, the whiteboard, the black drawings on the whiteboard, and the helper's hands were all considered foreground objects and hence became the shared view. Our observations indicated that since the helper was more knowledgeable about trigonometry and had access to formulas and a calculator, it was useful for the worker to observe the helper solve the problem in the helper's own space and then replicate the solution. We believe this idea can be successfully applied to industrial assembly scenarios, like the "tool"-scenario, if the helper is equipped with a 1:1 or miniature model of the machinery/tool on-site. We leave this as an interesting future direction of research to investigate the usefulness of manipulating real physical objects or replica scale models in comparison to using virtual object models [41,42] for remote assistance.

5.4. Camera work

We define camera work as the effort by the worker to move and position the camera with the purpose of communicating with the helper and regard it as a specialization of tailoring work.

Interestingly, we observed a pattern across scenarios, where workers typically used a scene camera most of the time (i.e. it was mounted in the environment), because it offered hands-free use, but occasionally picked up the camera in one hand for short periods of time to provide a close-up view of an object or move the camera to another location. This pattern was used occasionally with a webcam on a gorillapod in the "trigo" and "tool"-scenarios, but was particularly prevalent in the "injection"-scenario, where workers moved the tablet camera on a tripod back and forth between human machine interface and window to machine internals (see Fig. 9, b and c). However, this use of *movable scene cameras* led to a challenge in the "trigo"-scenario during the use of a projector in combination with a webcam on a gorillapod. A worker wanted to pick up the webcam and move it closer to the whiteboard to provide a closer view of some calculations. However, he was prevented from doing so, because that would require a re-calibration of the projection mapping. He voiced his frustrations to the experimenters afterwards stating that he preferred another configuration without the projector because of this issue. We discuss this limitation of the projection component in Section 6.

Another pattern observed in the "injection"-scenario was that workers often needed to show a very detailed view of the manufactured plastic pieces to helpers so they could evaluate the quality of the pieces. Workers would therefore move a piece close to the camera – about 5 cm from the lens – and rotate it in all directions for their helpers to get a good view of the imperfections in the piece (see Fig. 9, a). In the post-study questionnaire, one of the participants explicitly expressed the need for a high quality video feed for remote quality control.

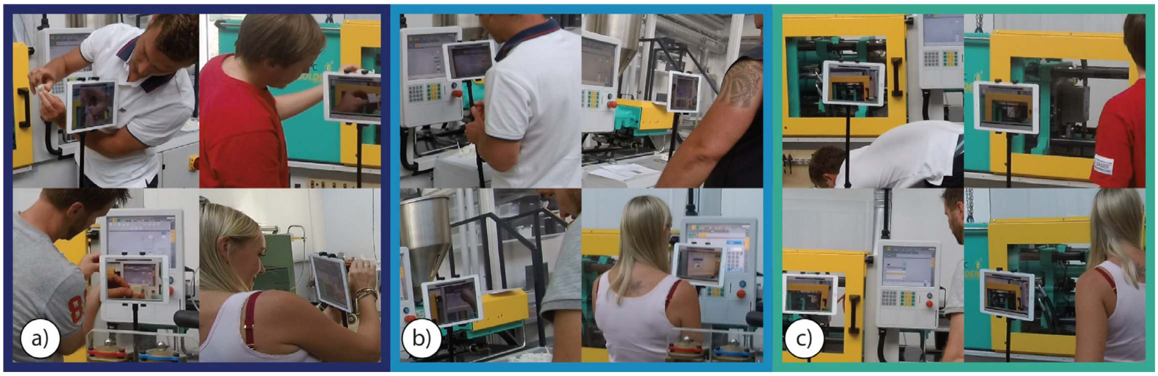

Fig. 9. Tailoring patterns in "injection"-scenario. (a) Workers provide close-up view of plastic pieces with imperfections. (b) Workers capture human machine interface. (c) Workers capture window to machine internals.

One of the workers, who used the webcam on a gorillapod in the "tool"-scenario, unexpectedly curled the legs of the gorillapod around his neck and left shoulder, thus providing a shoulder view from his perspective (see Fig. 8, b). Our observations show that this setup provided the helper with a stable view of the worker's task space in terms of being able to provide guidance with hands gestures. Another worker found a cardboard box, placed it on the workbench and put the camera on top of it to capture an overview of the available tools and mold (see Fig. 8, c). These are good examples of how technology is tailored in unexpected ways by the users.

5.5. Findings specific to the manufacturing context

Sharing of sound: We found evidence during the workshop with RAK that sound is regarded as an important source of information during remote troubleshooting of a running machine in the production. The main responsible teacher at the craftsmen school explained how important it is to listen to the machines to get information about the injection molding process. More specifically, he listens to process sounds on two sides of the machine: the "clamping side" with the shape forming mold that outputs plastic pieces and the "injection side", where the raw material is input, heated up, plasticized and transported to the mold. According to him, there is a rhythm to the process and rhythms that sound off can indicate various problems in the process that need to be addressed. He further emphasized that *"it will be difficult for a remote helper to give guidance unless he can hear the process sounds and see the process at the same time, because you use all senses in troubleshooting mode"*. He also considered other senses, as he perceived it as useful to be able to distinguish between the necessary smell of heated plastic from the production process and the erroneous smell of burned plastic from a control cabinet. Future research directions for the sharing of task space sounds are considered in the next section (see Section 6).

Multi-camera setup on machines: As previously mentioned, we observed how workers moved the camera back and forth between human machine interface and window to machine internals in the "injection"-scenario. A problem occurred, when a worker forgot to do so, as it was clear from one of the helper's utterances that she lost awareness of the worker's whereabouts, when the worker moved to another location on the machine without bringing the camera and verbally articulating his transition. In agreement with this observation, the same helper expressed the need for simultaneous views of the human–machine interface and machine internals of the injection molding machine because the two areas are connected, i.e. changes to parameters on the interface affects how the machine internals behave, which points to the usefulness of using multiple movable scene cameras to maintain workspace awareness. Machines usually have at least three distinct areas of interest during troubleshooting, the "injection side", "clamping side" and human machine interface. Multiple webcams are already supported in RAK, but the helper was unfortunately not aware of this.

6. Discussion

6.1. Design implications

We have derived four design requirements based on the discussion of the related work for our tailorable RAK system. These include:

1. RAK supports components for changing the guidance format (drawings or hands) and guidance location (on external display or projected into the task space).
2. Through the task space capturing components, RAK supports both view dependence and view independence, with the webcam(s) offering a view independent option and the smartphone/tablet offering both a view dependent and independent option, depending on whether the smartphone/tablet is handheld or mounted in the environment.
3. RAK implements negotiability and supports negotiation during the composition of configurations.
4. RAK supports projected guidance and provides different guidance formats and various camera capturing methods.

Our user study with RAK has generated findings that are specific to remote guidance systems. We discuss their possible design implications as below.

One design implication for the finding that users tailor remote assistance, including task space capturing and guidance format, to the distinct properties of scenarios, is the idea of recommended configurations, i.e. based on a given scenario RAK recommends a specific configuration proven to be useful in the scenario. For instance, given a scenario that entails troubleshooting a human–machine interface, RAK could remind the worker that using a configuration with a combination of overview and detailed cameras is likely to improve a remote helper's workspace awareness.

An implication for the finding that awareness of the collaborating partner's composition of components can be important, exemplified by the importance of the helper's mental model of guidance location, is the explicit design of persistent awareness cues in the remote assistance interface letting users know of their collaborating partner's configuration and notifying them about changes to the configuration during run-time. In the case of RAK, this means that a helper must be made aware of the worker's choices for task space capturing and additional guidance location choices in the Helper App's interface. Vice versa, the worker must be made aware of the helper's choices for interface mobility and guidance format in the Worker App and Supplementary Worker App.

The finding that many workers would move a scene camera around a workspace to cover different areas and sometimes pick it up to zoom into areas has implications for the design of remote assistance with augmented reality (AR). Advanced state-of-the-art remote assistance typically uses AR based head-mounted displays [43,44], and the workspace is thus captured from the worker's field of view and the helper's guidance is given from this perspective. Alternatively, we propose to study the combined benefits of movable scene cameras and head-mounted AR, where the helper's guidance originates from a movable scene camera, but is perceived by the worker on an AR-HMD. We know of two systems that combine movable scene cameras and head-mounted AR [45,46], but their research focus was not on industrial applications. We propose to study the benefits and drawbacks of combining movable scene cameras with head-mounted AR in industrial applications such as remote quality control, where close-up views are particularly important to a helper.

The finding that machine sounds should be shared in real world manufacturing factories is surprising and we discuss its implications for system design in the next subsection.

6.2. Sharing of sounds from the task environment

It was considered important for the remote helper to clearly hear the sounds from the running injection molding machine, as this enabled them to gather critical information about its state. This shows that the sharing of audio from the worker's task space, including not only the voice of the worker but also the sounds from the machinery, is important to the helper's situational awareness in some cases. However, most research on remote assistance has predominantly focused on visual aspects of a task space sharing, such as different methods of achieving view independence [8,20,22,23,25,47,48] and focus-in-context views [46,49,50]. Research-based remote assistance systems typically implement audio communication as an afterthought, because it is not the focus of the study, and in some lab studies the system implementation of audio sharing is not even needed, because the worker and helper participants are in the same room and can naturally hear each other.

In remote assistance scenarios in the manufacturing industry, such as the "injection"-scenario, it seems that the concept of *machine noise enhancement* can be used to support the collaboration between a worker and a helper. Thus, we believe that careful design of how task space audio is shared and enhanced is an interesting and industrially relevant area for future research. For example, the concepts of interactive noise cancellation and noise enhancement can be implemented using remote assistance in mixed reality. Multiple microphones could be distributed in an industrial environment. A remote helper in virtual reality points to or approaches a sound-generating task space object in the environment (this could be the "closed side" or the "injection side" of an injection molding machine), whereby the microphone closest to the object turns on, while the remaining microphones turn off. When the helper turns his attention to the worker in augmented reality by looking at him, the worker's wearable microphone turns on instead of the microphones in the environment. As a result, noise from machinery or clear verbal communication with the worker can be interactively selected.

6.3. Limitations and future work

We had 12 domain experts (6 pairs, N = 12) participate in our experimental simulation. This limited number of participants weakens our analysis of the quantitative data and has prevented us from conducting a more thorough statistical analysis of the results. Therefore, we cannot obtain good generalizability of the observed patterns in tailoring work. Instead, we have focused our analysis and discussion more on qualitative information. On the other hand, the strength of our data lies in the realism of the scenarios within which data about tailorable remote assistance was gathered and the participants' average of 11 years of experience with plastic manufacturing. This led to instances of unexpected tailoring work and revealed some interesting design requirements for remote assistance in the manufacturing industry.

Regarding future improvements of the RAK system, some helpers found it challenging to draw at the small scale required when workers used projected guidance. This usability problem was caused by a combination of the limited screen real estate of the helpers' tablet and the fixed scene camera view needed for projecting guidance correctly onto a planar surface in the workers' space. However, the projector component can be improved by using the camera, which is normally used for task space capturing, to track the pose of the planar surface. This would cause the homography needed for correctly warping a helper's guidance to be continuously recalculated, ensuring correct projection of guidance when the camera is moved, as demonstrated by Adcock et al. [51]. With this improvement, a worker will be able to pick up a camera and move it closer to a planar surface, thus supporting movable scene cameras, which was a common pattern in camera work, and making it easier for the helper to draw details on the surface.

As for future evaluation with RAK, the next step is a longitudinal field study to observe how its configurations are used in the manufacturing industry. In this paper, we collected data on how configurations were used after a short period. We expect new insights from a field study, as the use of RAK is likely to evolve with users' increased mastery and understanding of the strengths and drawbacks of its configurations in different scenarios.

7. Conclusion

We conducted an experimental simulation with a tailorable remote assistance system, RAK, involving professionals from the plastic manufacturing industry. Our findings revealed several insights related to tailorable remote assistance and its requirements in the manufacturing sector: (1) Users tailor remote assistance to the properties of problem scenarios and perceive this as useful. (2) Movable scene cameras were a dominant camera work pattern in the "injection"-scenario due to the need for close-up views of different machine areas. (3) Awareness of the collaborating partner's tailoring work is crucial. This was exemplified by the helper's lack of awareness of the guidance location component and its effect on the intelligibility of guidance in the worker's space. (4) Sharing machine sounds is important for remote troubleshooting of a running machine. (5) A multi-camera setup at manufacturing machines may improve awareness.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The author is an Editorial Board Member/Editor-in-Chief/Associate Editor/Guest Editor for Computers & Graphics and was not involved in the editorial review or the decision to publish this article.

References

[1] von Hippel Eric. Democratizing innovation: The evolving phenomenon of user innovation. J Betriebswirtsch 2005;55(1):63–78.
[2] Huang Weidong, Billinghurst Mark, Alem Leila, Xiao Chun, Rasmussen Troels. Computer-supported collaboration: theory and practice. John Wiley & Sons; 2024.
[3] Rasmussen Troels Ammitsbøl, Gronbak Kaj. Tailorable remote assistance with RemoteAssistKit: A study of and design response to remote assistance in the manufacturing industry. In: Collaboration technologies and social computing: 25th international conference, cRIWG+ collabTech 2019, kyoto, Japan, September 4–6, 2019, proceedings 25. Springer; 2019, p. 80–95.
[4] McGrath Joseph E. Methodology matters | Human-computer interaction.
[5] Stevenson Duncan, Li Jane, Smith Jocelyn, Hutchins Matthew. A collaborative guidance case study. In: Plimmer Beryl, Weber Gerald, editors. Ninth australasian user interface conference. AUIC 2008, Vol. 76, ACS; p. 33–42.
[6] Domova Veronika, Vartiainen Elina, Englund Marcus. Designing a remote video collaboration system for industrial settings. In: Proceedings of the ninth ACM international conference on interactive tabletops and surfaces - ITS '14. Dresden, Germany: ACM Press; 2014, p. 229–38.

[7] Speicher Maximilian, Cao Jingchen, Yu Ao, Zhang Haihua, Nebeling Michael. 360Anywhere: Mobile ad-hoc collaboration in any environment using 360 video and augmented reality. Proc ACM Hum-Comput Interact 2018;2(EICS):1–20.

[8] Tait Matthew, Billinghurst Mark. The effect of view independence in a collaborative AR system. Comput Support Coop Work (CSCW) 2015;1–27.

[9] Kraut Robert E, Miller Mark D, Siegel Jane. Collaboration in performance of physical tasks: effects on outcomes and communication. In: Proceedings of the 1996 ACM conference on computer supported cooperative work. CSCW '96, New York, NY, USA: Association for Computing Machinery; 1996, p. 57–66.

[10] Fussell Susan R, Setlock Leslie D, Kraut Robert E. Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In: Proceedings of the SIGCHI conference on human factors in computing systems. CHI '03, New York, NY, USA: Association for Computing Machinery; 2003, p. 513–20.

[11] Fussell Susan R, Setlock Leslie D, Yang Jie, Ou Jiazhi, Mauer Elizabeth, Kramer Adam DI. Gestures over video streams to support remote collaboration on physical tasks. Hum-Comput Interact 2004;19(3):273–309.

[12] Huang Weidong, Alem Leila. Gesturing in the air: Supporting full mobility in remote collaboration on physical tasks. J. Univers. Comput. Sci. 2013;19(8):1158–74.

[13] Huang Weidong, Kim Seungwon, Billinghurst Mark, Alem Leila. Sharing hand gesture and sketch cues in remote collaboration. J Vis Commun Image Represent 2018.

[14] Kirk David, Stanton Fraser Danae. Comparing remote gesture technologies for supporting collaborative physical tasks. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM; p. 1191–200.

[15] Kim Seungwon, Lee Gun A, Sakata Nobuchika. Comparing pointing and drawing for remote collaboration. In: 2013 IEEE international symposium on mixed and augmented reality. ISMAR, 2013, p. 1–6.

[16] Li Jane, Wessels Anja, Alem Leila, Stitzlein Cara. Exploring Interface with Representation of Gesture for Remote Collaboration. In: Proceedings of the 19th australasian conference on computer-human interaction: entertaining user interfaces. ACM; p. 179–82.

[17] Gauglitz Steffen, Nuernberger Benjamin, Turk Matthew, Höllerer Tobias. World-stabilized annotations and virtual scene navigation for remote collaboration. In: Proceedings of the 27th annual ACM symposium on user interface software and technology - UIST '14. Honolulu, Hawaii, USA: ACM Press; 2014, p. 449–59.

[18] Fakourfar Omid, Ta Kevin, Tang Richard, Bateman Scott, Tang Anthony. Stabilized Annotations for Mobile Remote Assistance. In: Proceedings of the 2016 CHI conference on human factors in computing systems. ACM; p. 1548–60.

[19] Wickey Aiden, Alem Leila. Analysis of hand gestures in remote collaboration: some design recommendations. In: Proceedings of the 19th australasian conference on computer-human interaction: entertaining user interfaces. ACM; p. 87–93.

[20] Fussell Susan R, Setlock Leslie D, Kraut Robert E. Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM; p. 513–20.

[21] Kurata Takeshi, Sakata Nobuchika, Kourogi Masakatsu, Kuzuoka Hideaki, Billinghurst Mark. Remote collaboration using a shoulder-worn active camera/laser. In: Proceedings of the eighth international symposium on wearable computers. ISWC '04, USA: IEEE Computer Society; 2004, p. 62–9.

[22] Lanir Joel, Stone Ran, Cohen Benjamin, Gurevich Pavel. Ownership and Control of Point of View in Remote Assistance. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM; p. 2243–52.

[23] Adcock Matt, Anderson Stuart, Thomas Bruce. RemoteFusion: real time depth camera fusion for remote collaboration on physical tasks. In: Proceedings of the 12th ACM SIGGRAPH international conference on virtual-reality continuum and its applications in industry - VRCAI '13. Hong Kong, Hong Kong: ACM Press; 2013, p. 235–42.

[24] Kim Seungwon, Billinghurst Mark, Lee Gun. The effect of collaboration styles and view independence on video-mediated remote collaboration. Comput Support Coop Work (CSCW) 2018;27(3):569–607.

[25] Lee Gun A, Teo Theophilus, Kim Seungwon, Billinghurst Mark. A user study on MR remote collaboration using live 360 video. In: 2018 IEEE international symposium on mixed and augmented reality. ISMAR, 2018, p. 153–64, ISSN: 1554-7868.

[26] Teege Gunnar. Users as Composers: Parts and Features as a Basis for Tailorability in CSCW Systems. Comput Support Coop Work (CSCW) 2000;9(1):101–22.

[27] Mørch Anders. Three levels of end-user tailoring: customization, integration, and extension. In: Computers and design in context. Cambridge, MA, USA: MIT Press; 1997, p. 51–76.

[28] MacLean Allan, Carter Kathleen, Lövstrand Lennart, Moran Thomas. User-tailorable systems: pressing the issues with buttons. In: Proceedings of the SIGCHI conference on human factors in computing systems. CHI '90, New York, NY, USA: Association for Computing Machinery; 1990, p. 175–82.

[29] Kaplan Simon M, Carroll Alan M. Supporting collaborative processes with Conversation Builder. Comput Commun 1992;15(8):489–501.

[30] Stiemerling O, Cremers AB. Tailorable component architectures for CSCW-systems. In: Proceedings of the sixth euromicro workshop on parallel and distributed processing - PDP '98 -. 1998, p. 302–8.

[31] Herrmann Thomas. Workflow management systems: ensuring organizational flexibility by possibilities of adaptation and negotiation. In: Proceedings of conference on organizational computing systems. COCS '95, New York, NY, USA: Association for Computing Machinery; 1995, p. 83–94.

[32] O'Neill Jacki, Castellani Stefania, Roulland Frederic, Hairon Nicolas, Juliano Cornell, Dai Liwei. From ethnographic study to mixed reality: a remote collaborative troubleshooting system. In: Proceedings of the ACM 2011 conference on computer supported cooperative work. ACM; p. 225–34.

[33] Palmer Doug, Adcock Matt, Smith Jocelyn, Hutchins Matthew, Gunn Chris, Stevenson Duncan, Taylor Ken. Annotating with light for remote guidance. In: Proceedings of the 19th australasian conference on computer-human interaction: entertaining user interfaces. ACM; p. 103–10.

[34] Alem Leila, Huang Weidong, Tecchia Franco. Supporting the changing roles of maintenance operators in mining: A human factors perspective. Ergon Open J 2011;4(1).

[35] Lukosch Stephan, Lukosch Heide, Datcu Dragoş, Cidota Marina. Providing information on the spot: Using augmented reality for situational awareness in the security domain. Comput Support Coop Work (CSCW) 2015;24(6):613–64.

[36] von Hippel Eric, Katz Ralph. Shifting innovation to users via toolkits. Manage Sci 2002;48(7):821–33.

[37] Franke Nikolaus, Hippel Eric von. Satisfying heterogeneous user needs via innovation toolkits: the case of Apache security software. In: Open source software development, Res Policy In: Open source software development, 2003;32(7).1199–215,

[38] Huang Weidong, Alem Leila. HandsinAir: a wearable system for remote collaboration on physical tasks. In: Proceedings of the 2013 conference on computer supported cooperative work companion. CSCW '13, San Antonio, Texas, USA: Association for Computing Machinery; 2013, p. 153–6.

[39] Clark Herbert H, Brennan Susan E. Grounding in communication. In: Perspectives on socially shared cognition. Washington, DC, US: American Psychological Association; 1991, p. 127–49.

[40] Kraut Robert E, Gergle Darren, Fussell Susan R. The use of visual information in shared visual spaces: Informing the development of virtual co-presence. In: Proceedings of the 2002 ACM conference on computer supported cooperative work. ACM; p. 31–40.

[41] Oda Ohan, Feiner Steven. 3D referencing techniques for physical objects in shared augmented reality. In: 2012 IEEE international symposium on mixed and augmented reality. ISMAR, Atlanta, GA, USA: IEEE; 2012, p. 207–15.

[42] Adcock Matt, Ranatunga Dulitha, Smith Ross, Thomas Bruce H. Object-based touch manipulation for remote guidance of physical tasks. In: Proceedings of the 2nd ACM symposium on spatial user interaction. SUI '14, Honolulu, Hawaii, USA: Association for Computing Machinery; 2014, p. 113–22.

[43] Lee Gun A, Teo Theophilus, Kim Seungwon, Billinghurst Mark. Mixed reality collaboration through sharing a live panorama. In: SIGGRAPH Asia 2017 mobile graphics & interactive applications. 2017, p. 1–4.

[44] Teo Theophilus, Lee Gun A, Billinghurst Mark, Adcock Matt. Hand gestures and visual annotation in live 360 panorama-based mixed reality remote collaboration. In: Proceedings of the 30th Australian conference on computer-human interaction. OzCHI '18, Melbourne, Australia: Association for Computing Machinery; 2018, p. 406–10.

[45] Piumsomboon Thammathip, Lee Gun A, Irlitti Andrew, Ens Barrett, Thomas Bruce H, Billinghurst Mark. On the shoulder of the giant: A multi-scale mixed reality collaboration with 360 video sharing and tangible interaction. In: Proceedings of the 2019 CHI conference on human factors in computing systems. CHI '19, Glasgow, Scotland Uk: Association for Computing Machinery; 2019, p. 1–17.

[46] Rasmussen Troels Ammitsbøl, Huang Weidong. SceneCam: Improving multi-camera remote collaboration using augmented reality. In: 2019 IEEE international symposium on mixed and augmented reality adjunct. ISMAR-adjunct, 2019, p. 28–33, ISSN: null.

[47] Kuzuoka Hideaki, Oyama Shinya, Yamazaki Keiichi, Suzuki Kenji, Mitsuishi Mamoru. GestureMan: a mobile robot that embodies a remote instructor's actions. In: Proceedings of the 2000 ACM conference on computer supported cooperative work - CSCW '00. Philadelphia, Pennsylvania, United States: ACM Press; 2000, p. 155–62.

[48] Ranjan Abhishek, Birnholtz Jeremy P, Balakrishnan Ravin. An exploratory analysis of partner action and camera control in a video-mediated collaborative task. In: Proceedings of the 2006 20th anniversary conference on computer supported cooperative work. ACM; p. 403–12.

[49] Yamaashi Kimiya, Cooperstock Jeremy R, Narine Tracy, Buxton William. Beating the limitations of camera-monitor mediated telepresence with extra eyes. In: Proceedings of the SIGCHI conference on human factors in computing systems. CHI '96, New York, NY, USA: ACM; 1996, p. 50–7.

[50] Norris James, Schnädelbach Holger, Qiu Guoping. CamBlend: An object focused collaboration tool. In: Proceedings of the SIGCHI conference on human factors in computing systems. CHI '12, New York, NY, USA: ACM; 2012, p. 627–36.

[51] Adcock Matt, Gunn Chris. Using Projected Light for Mobile Remote Guidance. Comput Support Coop Work (CSCW) 2015;24(6):591–611.