Authorable Augmented Reality Instructions for Assistance and Training in Work Environments

Jonas Blattgerste
University of Applied Sciences
Emden / Leer
Emden, Germany
jblattgerste@techfak.uni-bielefeld.de

Patrick Renner
CITEC - Cluster of Excellence
Cognitive Interaction Technology
Bielefeld, Germany
prenner@techfak.uni-bielefeld.de

Thies Pfeiffer
University of Applied Sciences
Emden / Leer
Emden, Germany
thies.pfeiffer@hs-emden-leer.de

ABSTRACT

Augmented Reality (AR) is a promising technology for assistance and training in work environments, as it can provide instructions and feedback contextualised. Not only, but especially impaired workers can benefit from this technology. While previous work mostly focused on using AR to assist or train specific predefined tasks, "general purpose" AR applications, that can be used to intuitively author new tasks at run-time, are widely missing.

The contribution of this work is twofold: First we develop an AR authoring tool on the Microsoft HoloLens in combination with a Smartphone as an additional controller following considerations based on related work, guidelines and focus group interviews. Then, we evaluate the usability of the authoring tool itself and the produced AR instructions on a qualitative level in realistic scenarios and gather feedback. As the results reveal a positive reception, we discuss authorable AR as a viable form of AR assistance or training in work environments.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI)**.

KEYWORDS

Augmented Reality, Mixed Reality, Authoring, Annotation, Assistance, Training, Cognitive impairments

ACM Reference Format:
Jonas Blattgerste, Patrick Renner, and Thies Pfeiffer. 2019. Authorable Augmented Reality Instructions for Assistance and Training in Work Environments. In *MUM 2019: 18th International Conference on Mobile and Ubiquitous Multimedia (MUM 2019), November 26–29, 2019, Pisa, Italy*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3365610.3365646

1 INTRODUCTION

In work environments, Augmented Reality (AR) is a viable means of displaying instructions and feedback contextualised to assist or train new workers. That AR outperforms conventional instructions in assistance scenarios has already been shown in multiple studies [7, 10, 43] and studies on AR training show clear indications for improved and simplified learning [25, 44]. Beside workers that train on a new task, especially workers with cognitive impairments could also greatly benefit from such assistance systems in work environments [3, 21]. Many of them are illiterate, which makes conventional instructions challenging [17], and learning from experts directly is often cognitively demanding [26].

Due to the increasing attention and technological advancements (e.g. ARCore/ARKit or the Microsoft HoloLens), research on AR using handheld and head-mounted devices is currently transitioning from the "technical feasibility" research phase, where experiments are done in controlled laboratory settings, to evaluating AR in real scenarios and with larger sample sizes [11].

This transition is certainly desirable for the research field, but reveals a gap in the current body of literature. Looking at secondary literature on AR instructions in work environments, it becomes apparent that right now, the vast majority of applications assisting through AR instructions are developed for predefined tasks [5, 30, 42] and do not incorporate functionality to easily create new AR instructions. This issue has also already been discussed for maintenance, e.g., by Zhu et al. [49]. This approach is intentional in the "technical feasibility" research phase to answer specific questions and to get an understanding of the general reception of AR in these settings. However, the transition into evaluating AR in realistic work environments provokes the question of the viability and usability of AR instructions that are not specifically developed for the task they are evaluated in (see Figure 1, right). We believe that for the vast majority of tasks in work environments, specialised AR applications are not mandatory, but rather modular "general purpose" applications would be required which can be used to intuitively and quickly create new AR instructions after being deployed (see Figure 1, left). This could be especially beneficial for workers with cognitive impairments. As the individual impairments usually differ significantly from person to person, at the same time the needed kind of assistance differs. Using a flexible authoring tool, individually tailored solutions could easily be created.

To contribute towards answering the question of the viability and usability of authorable AR instructions in work environments, the contribution of this paper is twofold. First, we develop and implement an AR authoring tool for the Microsoft HoloLens, drawing upon related work, design guidelines, and focus group interviews with workers and instructors the application is targeted at. Secondly, we evaluate the authoring and the usage of authored AR instructions using our developed authoring tool on a primarily qualitative level and discuss the results and implications.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
MUM 2019, November 26–29, 2019, Pisa, Italy
© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-7624-2/19/11…$15.00
https://doi.org/10.1145/3365610.3365646

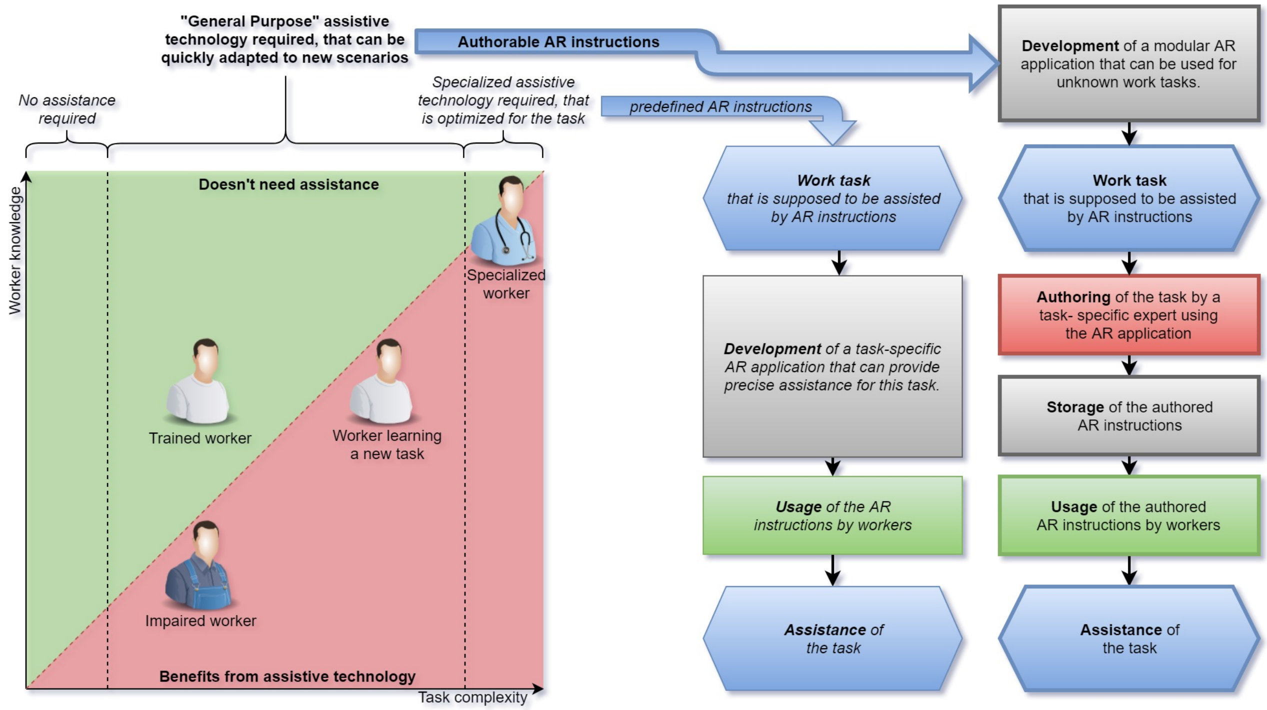

Figure 1: The relationship of the complexity of a task and the preexisting task knowledge of the worker (left) with its implications on when and which type of AR assistance would be best suited (right). While for specialised work tasks, predefined AR instructions are necessary, for the majority of tasks, AR instructions authored through modular AR tools should be sufficient.

2 RELATED WORK

While the early endeavours of creating tools for the authoring of new AR content date back to almost 20 years ago with publications by Haringer et al. [18] or Poupyrev et al. [32], both describing AR authoring systems for a fast creation of AR content using markers, further research on intuitive AR authoring tools has only been sparse compared to the general expansion of the research field. Furthermore, publications on authoring tools often only describe a framework or a prototype but do not evaluate the proposed tool. Both research gaps could at least partially be explained by the technical feasibility of authoring tools. In general, AR instructions until recently (before the release of ARCore, ARKit or HoloLens) suffered from insufficient tracking capabilities of the available hardware and made realistic AR usage challenging. Although sparsely covered and rarely directly applicable to work environments, some noteworthy primary research on AR authoring tools was conducted over the years. Generally those endeavours can be split into two categories: integrated authoring capabilities of the application (the AR application itself is used to author new AR instructions) and external authoring tools (desktop-based tools or scripting languages are used to create new instructions for the AR device).

2.1 External Authoring Tools

One notable external authoring tool is the designers augmented reality toolkit (DART) for quick AR prototyping proposed by MacIntyre et al. [27]. They use "Adobe Director" in combination with a scripting language to combine predefined AR specific components to rapidly prototype interactive AR experiences on both handheld and head-mounted devices. Expanding this research, Seichter et al. [37] designed a standalone application to intuitively create interactive AR content based on DART, still based on scripting with predefined components. Similarly, Zhu et al. [49] developed a desktop-based authoring interface to author AR instructions for head-mounted displays (HMDs) in maintenance tasks. Their desktop interface also enables users with low-level programming skills and a mouse to author AR instructions, similar to what game engines like Unity offer. Finally, Zauner et al. [48] proposed an AR assistance application for hierarchical assembly tasks (e.g. building furniture), including the possibility to create or edit the assembly instructions through editing an XML file, that is then used on an AR HMD to display the assembly instructions for the task.

2.2 Integrated Authoring Capabilties

Regarding integrated authoring approaches, Piekarski et al. [31] proposed a glove-based AR authoring system that could be used to create simple three-dimensional shapes for architectural modelling by visually tracking position and gesture of the user's hands. Similarly, an application for in-situ AR content creation in unprepared outdoor environments on handheld devices was proposed by Langhtz et al. [22]. In their proposed application, instructions are additionally stored on a server to be easily accessed by others.

Guven et al. [16] developed a handheld AR authoring tool, utilising touch interaction and computer-vision, to place AR content (e.g. hypermedia or multimedia) into the physical environment.

Funk et al. [13, 14] developed a stationary projection-based AR authoring tool that automatically detects and authors AR instructions based on movement detection within the assembly area using a Microsoft Kinect. They compare their approach to manual programming the AR instructions and show that their approach is significantly faster but similarly cognitively demanding. In line with their approach, Lee et al. [24] developed an AR authoring tool for stationary AR systems. They use markers and marker-cubes as tangible objects to create, position, and interact with new AR content.

In the context of people with cognitive impairments, Quintana et al. [33] developed an application aimed at the elderly with Alzheimers and their caregivers which can be used to author tasks or reminders for the elderly on a handheld device by combining computer vision with touch interaction. Similarly, Wolf et al. [46] developed an AR application for HMDs which can be used to assist elderly people with Alzheimers in daily life activities (e.g. cooking). Their application combines in-situ authoring and external authoring tools, as instructions are first created on a desktop interface and then loaded into the HMD and eventually placed onto the desired location utilising the HoloLens head-tracking and gesture interaction.

3 FOCUS GROUP INTERVIEWS

Before developing the actual application (Section 4), we evaluated an early mock-up prototype to get preliminary insights. The mock-up consisted of two sets of visual storyboards, one for the authoring and one for the usage phase, showing the process of authoring and using augmented reality instructions as envisioned by us. The figures consisted of photos of people with the HoloLens and a smartphone as an additional controller and Augmented Reality content was placed in the photo using photo editing software.

Qualitative interviews were conducted in two distinctive settings. The first setting, a vocational scenario at a higher education institution where students learn practical medical skills, is called the "SkillsLab". The second was an assistance scenario for cognitive impaired workers where employees with impairments operate and maintain a kiosk and is part of the "Bethel prowerk". To represent the requirements for workers with cognitive impairments, we invited a focus group consisting of 5 instructors (3 female) for cognitively impaired workers aged between 33 and 59 (average = 48, sd = 11), as at this stage the cognitively impaired workers could not provide the necessary meta-reflections and projections of imagined use. As this group also already covers general aspects of instruction design, we focused on the recipients of the instructions with the second focus group, by inviting 7 students (6 female) from the vocational setting, aged between 27 and 37 (average = 31.57, sd = 5.38), that already completed their practical vocational training. The concept was received so positively in the focus group, that we were invited to conduct the practical evaluation of the prototype in the two organizations. This underlines the relevance the two organizations see in this work.

3.1 Study Design

In line with the guidelines for qualitative research methods proposed by Berg et al. [3] and practical strategies for combining qualitative and quantitative methods by Morgan et al. [28], our qualitative interviews were conducted as a complementary method in form of a preliminary study. As only an exploratory complement to our principal method, it is not supposed to be representative or conclusive.

We used standardised open-ended questions with the possibility to ask non-predefined followup questions to allow the participants to fully express their viewpoints and opinions [41]. An interview guide incorporating ten open-ended questions divided in two groups was used. The first part focused on the existing tasks in the scenario (e.g. what are common tasks, how are they currently assisted or trained). Then participants were shown and explained the storyboard showing our vision for the authoring tool. Afterwards, the second part focused on first impressions, feedback, suggestions and preferences regarding the shown mock-up.

3.2 Results & Implications

The resulting transcripts were coded according to Burnard et al. [9]. The most interesting common schemes and insights for the development of the authoring tool are as follows.

Both the instructors for cognitively impaired workers and the students in the vocational setting perceived our idea positively. The instructors stated that the primarily visual approach suits many of the impaired workers, as it is easy to understand and adds a motivation/fun factor because of the workers' interest in technical gadgets. Some additionally stated that being able to complete tasks independent of a human coaching by using the AR assistance could boost the worker's confidence. In line with this, the students stated that visual feedback and instructions are a desired learning modality and both could aid the training process, as often it is the many small considerations that make the learning process challenging, rather than the overall steps of a task. Noteworthy, some students stated that this approach might not be suitable or too challenging in training scenarios incorporating many social aspects (e.g. directly interacting with a patient) as wearing AR glasses could hinder the communication aspects.

Furthermore, the instructors stated that workers in their facilities can have a wide range of preexisting knowledge and capabilities. As a result, the approach the workers are trained and assisted with in their work environment is highly individual. Approaches vary in terms of necessary repetition, explanation speed and modality used. The instructors describe this as both a chance and a challenge for the application. On the one hand the ability to author different instructions for different people (e.g. some entirely based on pictures for illiterate workers and some mainly based on text instructions and arrows) is perceived as promising, on the other hand, it might be challenging to identify when which kind of instruction is best suited for the individual.

When asked how they think instructors and workers would receive such an application, both groups stated they think workers would generally receive the application positively but instructors might be hesitant because of their age and unfamiliarity with technology in general. They perceive the usability of the authoring tool

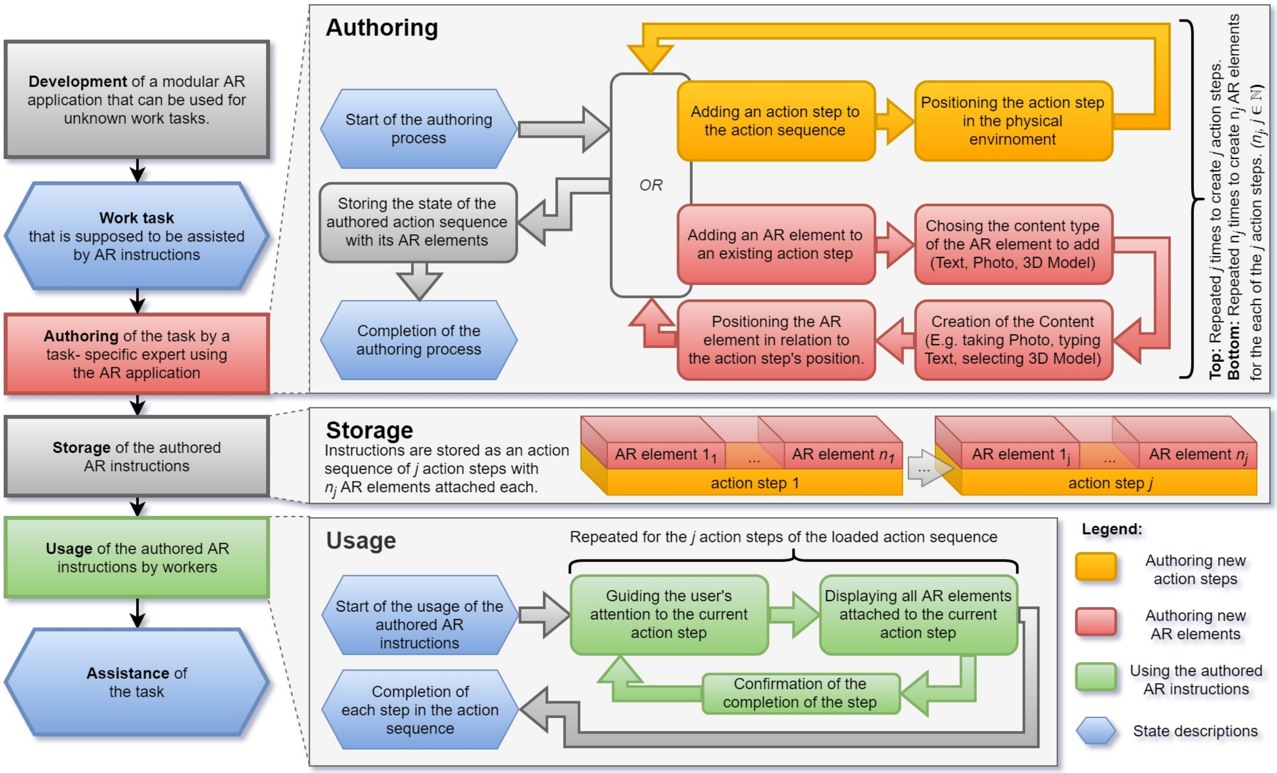

Figure 2: The processes of the Authoring, Storage and Usage of AR instructions on a conceptual level. The Authoring process incorporates the placement and positioning of action steps (orange) and their respective AR elements (red). AR instructions are then stored as a sequence of action steps with AR elements attached to each. Stored instructions can be loaded and used by the workers as a sequence of action steps that display their attached AR elements (green).

as the key factor. They stated that they believe that if the tool is not intuitive and self-explanatory, instructors might reject it entirely.

Finally, when asked what they would wish for in such a system, when it would be actually used in their work environment, both groups answered that most importantly the provided instructions have to be understandable for the worker. Additionally, some instructors said that a wide variety of available instruction modalities (e.g. pictures, text, models or videos) to choose from would be desirable and that they would wish for a continuous integration into the development process itself to provide continuous feedback.

4 THE AR AUTHORING TOOL

Combining our initial vision, the feedback from the focus groups and the implications from related work with additional sources on AR instructions and interaction, we developed an AR authoring tool on the Microsoft HoloLens in combination with a smartphone as an additional interaction device. At this point it should be underlined that our aim is to develop an authoring tool for a wide variety of possible assistance and training scenarios, ranging from people with cognitive impairments to healthy workers. However, the specific authored instructions for different groups of workers are left to the expertise of the instructors, i.e. our users. The HoloLens itself, in combination with the smartphone, acts as the AR authoring tool that can be used by instructors to author new AR instructions for specific work tasks. The instructions can be stored on the device itself. When using the HoloLens without the smartphone, the application can be used to display the authored AR instructions and subsequently be used by workers to be assisted in or trained at a specific work task (see Figure 2, left). Consequently, this tool would be categorised as AR authoring tool with integrated authoring capabilities (Section 2). We chose the integrated authoring approach to not only make the authoring of new instructions as intuitive and easy to use as possible (especially eliminating the requirement of understanding basic software development and programming to create new AR instructions many external tools have) but also to be independent from external devices and speed up the authoring process itself.

We implemented the application using the game engine Unity with one application running on the Microsoft HoloLens and the other one on the Android smartphone (Samsung Galaxy S7) as an additional controller device. While authoring, both applications are connected to a Wifi router to communicate through TCP/IP.

The smartphone controller can either be used as a handheld device or strapped to the wrist of the user. While the HoloLens itself has speech and especially gesture interaction that can be utilised, a smartphone as an additional input device provides several advantages. For example, it can be used to quickly type text, take photos, preview 3D models and comparable to a gaming "joystick controller" to position the AR instructions into the physical environment (see Figure 3, right). This helps with the adaption of the system, as the target user group (the instructors) are already familiar with interaction with smartphones. Additionally, previous work suggests that using the gesture interaction on the HoloLens for precise movements potentially creates discomfort for the user [47] and that gesture interaction on the HoloLens in general is often perceived as "difficult" [45].

When using the authored AR instructions, the HoloLens can be used independently of the smartphone, either by using the Microsoft Clicker or by using gesture interaction to navigate through the AR instructions. Additionally, for study purposes, in the usage phase the Hololens can also be controlled by a Bluetooth keyboard instead of gesture interactions or the Clicker.

The process of authoring, storing and eventually using the AR instructions for work tasks is described in the following subsections in detail and visualised as a conceptual overview in Figure 2. Additionally, a practical example of authoring a new AR instruction using the authoring tool is shown in Figure 3 and two practical examples of already authored instructions are shown in Figure 5.

4.1 Authoring new AR Instructions

After starting both applications on the HoloLens and the smartphone, the instructor has to link both devices in an initial step to connect them over a TCP/IP connection.

Then the application is ready for use and the instructor can start the authoring of new AR instructions for an action sequence. AR instructions for an action sequence in our application consist of a linear sequence of 1 to j ($j \in \mathbb{N}_{\neq}$) discrete action steps:

$$\mathbf{AS} = \{AS_1, ..., AS_j\}, j \ \in \mathbb{N}_{\neq}$$

Each of the action steps can have 1 to n_j ($n_j \in \mathbb{N}_{\neq}$) AR elements attached that are displayed simultaneously in the usage phase (see Figure 2). This distinction between the AR elements and the action step itself was chosen for several reasons. Primarily, this allows to easily use several different AR elements to describe one step of an action sequence for a task and therefore introduces a wide range of flexible options for the instructor as to how the tasks' AR instructions can be authored. Furthermore, it allows the instructors to use different interaction techniques, whether they need to cover low-precision actions for the general direction of an action step or high-precision actions for the exact position of AR elements. This combination of several interaction techniques can improve both accuracy and user experience in AR systems [29].

In our application, to place an action step into the physical environment, the instructor can press "Add action step" (see Figure 3, left) and place it into the environment by positioning it via head-gaze and pressing a confirmation button on the smartphone controller. Here, head-gaze allows to easily control the placement of action steps that do not require much accuracy. The instructor can place as many action steps as necessary.

After placing one or several action steps, the instructor can then also add AR elements to already placed action steps by pressing the "add element" button on the respective action step in the menu on the smartphone Controller (see Figure 3, left). AR elements are then attached to the action step and can be positioned in relation to it (see Figure 3). Again, the instructor can add as many AR elements as necessary to one action step. All AR elements of one step are later displayed simultaneously for that action step when using the authored AR instructions. Several different types of AR elements are available: Instructors can capture and place photo elements, type and place text elements or select and place 3D models from several existing model packages, such as basic shapes, workplace tools, arrows & warning signs or hand animations based on the most common hand movements in work environments (see Figure 4 and 5). This allows for the creation of AR instructions that cover the majority of requirements from previous demand analyses for AR assistance [45] and implications from our focus group interviews. Furthermore, it provides all necessary tools to create AR instructions that are in line with previously identified guidelines on AR instructions [38], best practices on designing effective step-by-step instructions [1] and the requirements and recommendations of ISO standard 9241-210 that focuses on the user centred design approaches of interactive systems.

4.2 Persisting AR Instructions

After the authoring of AR instructions is completed, they are stored as the linear sequence of j ($j \in \mathbb{N}_{\neq}$) discrete action steps, with their n_j ($n_j \in \mathbb{N}_{\neq}$) AR elements attached to them (see Figure 2, Storage).

On a technical level, this was implemented by combining game state serialisation (e.g. systematically saving 3D model names, file paths to pictures, inserted text, position and rotation for each of the action steps with their AR elements) and the *WorldAnchor* functionality of the HoloLens that allows to store and retrieve fixed anchor points in the physical environment to attach holographic AR content to it. This accomplishes that authoring, storage and usage of AR instructions do not require any external positioning support like marker-tracking, even between sessions and in multiple physical rooms.

4.3 Usage of Authored AR Instruction

When starting only the HoloLens application, it is in the usage mode and one of the stored action sequences can be loaded and used to assist or train a worker.

The worker can manually progress through all the action steps of the loaded action sequence using gesture interaction (by performing the "air tap" gesture) or using the Microsoft Clicker. Hereby, all the AR elements attached to the action step that is currently active are displayed simultaneously (see Figure 5). The AR elements are displayed in-situ at the position in the physical world where the authoring instructor placed them. The application hereby handles both the occlusion of overlapping AR elements and physical elements (e.g. furniture) in the room itself. Correct occlusion handling was previously identified to be a crucial factor for the unambiguous visualisation of AR instructions [6, 38]. The "anchor" of the action step itself (see Figure 3, right) is no longer visible in the usage mode.

If the currently displayed AR elements are not in the Field of View of the HoloLens and therefore are not immediately visible to the worker, the worker is guided towards the currently displayed

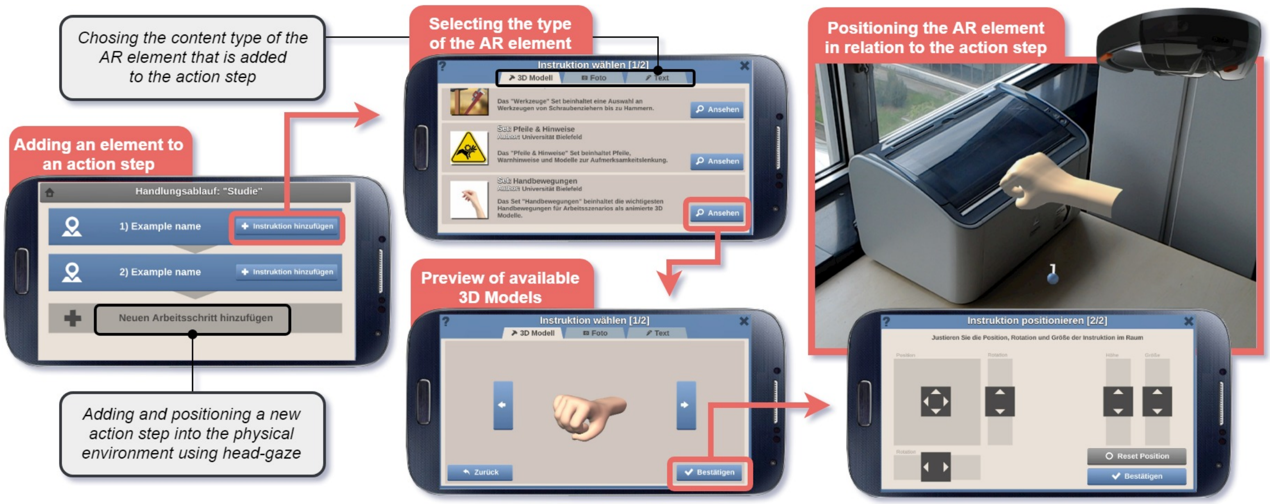

Figure 3: Practical example: The process of adding an AR element (hand animation) to an existing action step (1), that shows how to open the paper lid of a printer, using the authoring tool.

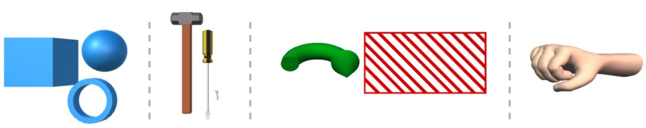

Figure 4: Actions can be depicted by 3D models: Basic shapes, tools, arrows & warning signs and hand animations.

AR elements. This is accomplished by using the SWave attention guiding technique proposed by Renner et al. [34].

5 EVALUATION: AR AUTHORING TOOL

First, we evaluated the usability of the AR authoring tool itself with instructors who would use the authoring tool to create AR instructions for assistance or training purposes in their work environments. Beside the usability and user experience, we additionally put an emphasis on qualitative feedback for the tool itself but also on their expectations of the acceptance and usability for the users of the authored instructions and the idea of "authoring AR instructions" beyond the proposed tool.

While comparative approaches (e.g. directly comparing authoring of AR instructions to the creation of traditional instructions in a laboratory setting) were considered, we ultimately decided that a usability evaluation with an emphasis on qualitative feedback from real instructors would provide most value, as it gives further insights into what their needs for such an application would be and the current state of the application is not a finished product.

5.1 Study Design & Procedure

The study was designed as a usability study. We used the System Usability Scale (SUS) [8] to get an easily comparable usability score, the User Experience questionnaire (UEQ) [23] to get more insights into the users experience and perception of the tool beyond the usability itself and a qualitative feedback questionnaire to enable the participants to voice their opinion freely. The qualitative feedback questionnaire focused on what they particularly liked or disliked, if they would use the authoring tool in their scenario and how they generally perceive the idea of authorable AR instructions for assistance and training purposes.

After explaining the experiment and given a short introduction to the authoring tool, participants were instructed to author the task that was identified to fit best in their work environment in the focus group interviews: The cleaning of a coffee maker in the assistance scenario for cognitively impaired workers and the preparation of an injection in the health-care vocational training scenario (see Figure 5). They were instructed that it is not necessary to fully complete the task but rather that they understand the tool and can provide feedback for it. After they had a good understanding of the application and used it to author parts of the respective task, they filled out the questionnaires.

5.2 Results

The application was tested and evaluated by 5 instructors from both evaluation scenarios. One from the health-care vocational training scenarios, 4 from the assistance scenario for cognitively impaired workers. They were aged between 28 and 65 (average = 48, sd = 14.3). Three of the instructors were female. The small sample size was caused by the unavailability of additional instructors.

5.2.1 Usability & User Experience. For usability, the instructors reported an average SUS score of 76 (sd = 12.6). According to Bangor et al. [2] this translates to a "good" usability, though with the sample size of 5 the SUS is not conclusive [40].

In terms of the reported user experience of the authoring tool, the average scores of the 6 measures captured by the UEQ were: 2,20 (sd = 0,32) for Attractiveness, 1,40 (sd = 1,05) for Perspicuity, 1,15 (sd = 0,72) for Efficiency, 1,40 (sd = 0,60) for Dependency, 2,25

Figure 5: One exemplary action step of the AR instructions, authored using the proposed authoring tool, for the evaluation in the assistance scenario for cognitively impaired workers (left) and in the health-care vocational training scenario (right).

(sd = 0,35) for Stimulation and 2,35 (sd = 0,52) for Novelty. Figure 6 contextualises theses scores into the global UEQ benchmark [35].

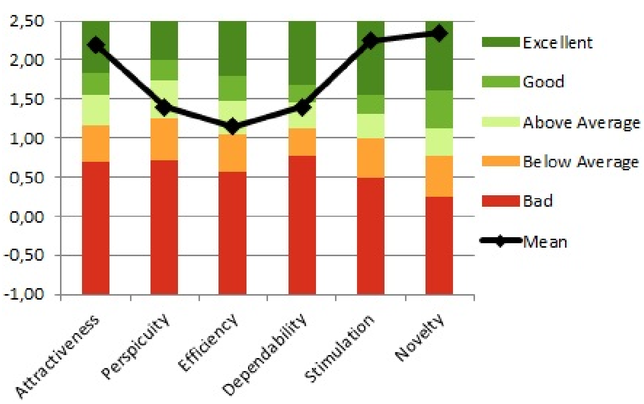

Figure 6: The User Experience Questionnaire (UEQ) scores reported by the instructors compared in the UEQ benchmark.

5.2.2 Qualitative Feedback. When asked what they particularly liked about the application, all instructors stated that they believe that the instructions they were able to create with the provided tools are easy to understand, especially describing them as "concise", "attractive", "simple" and "not overloaded". Furthermore, 2 instructors stated that they liked the availability of different options to chose from and another 2 stated that they liked the simple usability of the authoring tool. Additionally, 2 instructors particularly liked the innovation of the tool itself.

Facets of the authoring tool the instructors did not like were the hardware of the HoloLens, stating that the HoloLens is hard to wear with glasses (3 instructors), that the field of view was to narrow (2 instructors), and that performing precise adjustments with the smartphone controller "needs some practice time" (2 instructors).

All 5 instructors answered "yes", when asked if they would like to use such a tool for assistance/training purposes. As their reasons, they stated that they believe the resulting instructions can be created "very precise" and in an "easy to understand" manner (3 instructors), that users can work more independently and be "additionally motivated" by using instructions created by the authoring tool (3 instructors) and that the created AR instructions would be "received"/"accepted" well (2 instructors).

When asked what they think about the concept of letting instructors author new AR instructions through an authoring tool, all 5 instructors stated that they like the idea. Four instructors believed that instructors should know best what kind of instructions are needed for particular tasks or people, 2 instructors argued that the creation of new instructions is faster and one instructor stated that the primarily visual approach could be especially effective.

6 EVALUATION: AUTHORED AR INSTRUCTIONS

As authoring AR instructions is only viable when they are actually usable by the workers they are authored for, we furthermore evaluated acceptance and usability of instructions authored with the proposed tool in the two scenarios with real workers. The tasks used were not fabricated for the study and fit our scope of the application shown in Figure 1. They were identified in the focus group interviews and also used in the authoring evaluation (Section 5). The first is an assistance task for cognitively impaired workers, where the workers have to perform the cleaning of a modern coffee maker (see Figure 5) that incorporates 20 distinctive steps (e.g. pressing specific buttons and disassembling, emptying and cleaning specific parts of the machine). Note that in this study focusing on general usability and especially acceptance, the different specific impairments of the participants are not detailed. That this approach might be desirable for augmented reality based assistance systems was already discussed in [5]. The second task is a training task for health-care professionals, where students have to prepare an injection, incorporating a task with 20 distinctive steps (e.g. cleaning

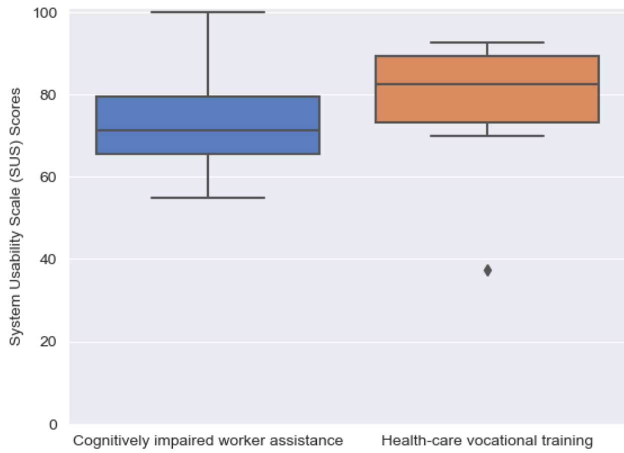

Figure 7: The System Usability (SUS) Scores for the authored AR instructions in both evaluation scenarios.

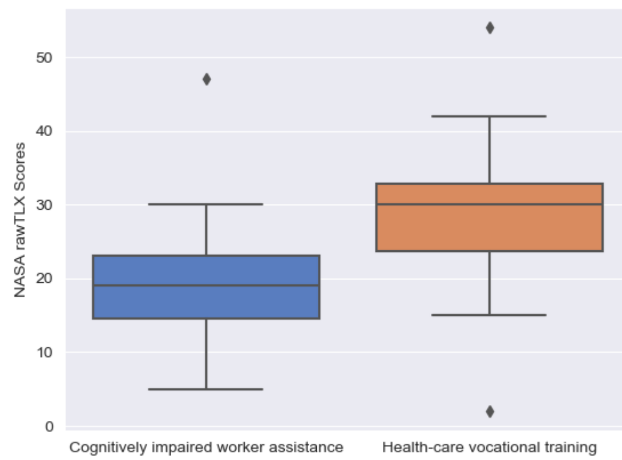

Figure 8: The perceived cognitive load (NASA rTLX) for the authored AR instructions in both evaluation scenarios.

the work space, sterilising hands, using disposable gloves, picking and assembling parts of the injection and venting the syringe).

6.1 Study Design

The studies for both scenarios were deliberately designed as pure usability studies without an A/B test to conventional instructions or the measurement of quantitative measures like task completion times (TCT). On the one hand this reduces stress factors and focuses the study on the important question, the usability, and on the other hand, for the assistance scenario of cognitively impaired workers, the instructors reported that none of the cognitively impaired workers were able to complete the task using the conventional instructions.

6.2 Procedure

Preparing the study, the experimenter and one of the instructors from the authoring tool evaluation authored the AR instructions for the two respective tasks.

During the experiment, first the participants were introduced to the experiment and the HoloLens. Then, they were explained what the upcoming AR instructions for the task aimed to accomplish. They were informed that they can ask for help if they do not understand the current instructions. No preliminary training sequence or task was used.

They completed the task using the AR instructions, while the experimenter controlled the application through Wizard of Oz and noted errors and help requests in a spreadsheet. Afterwards, participants were asked to fill out demographic, SUS and NASA TLX [19] questionnaires. Additionally, they were also asked for qualitative feedback and suggestions.

In the assistance scenario for cognitively impaired workers, participants were given the choice to read and write for themselves or get help by the experimenter. They were explained challenging words or sentences in easy language.

6.3 Results

Ten workers with differing cognitive impairments took part in the experiment of the assistance scenario for cognitively impaired workers. They were aged between 18 and 27 (average = 21.2, sd = 3.16), 5 were female.

Furthermore, 10 health-care students that were currently completing the practical phase of their study took part in the experiment of the vocational training for health-care professionals. They were aged between 19 and 35 (average = 23.3, sd = 5.03) and 7 of them were female.

6.3.1 Errors. All 20 participants were able to successfully complete the tasks using the AR instructions without major errors.

In the assistance scenario for cognitively impaired workers, 4 participants asked for additional help for the step of inserting a cleaning tablet on top of the machine, which was hard to reach from the front of the machine and 1 participant accidentally pressed a wrong button on the touchscreen of the machine during the process.

In the vocational training scenario, 2 participants had problems identifying the correct syringe, as several types were stored in the drawer and asked for additional help. Furthermore, 1 participant picked the wrong needle, as here again several types of needles were stored in the drawer. Another participant asked for additional help to identify the correct medication that is supposed to be picked for the injection. Furthermore, when putting on disposable gloves, this was recognised as the "bloom gesture" on the HoloLens for several participants. This opened the menu and paused the application. The participants were then shown how to close the menu again to proceed.

6.3.2 Usability. In the assistance scenario for cognitively impaired workers, the average reported SUS score was 74.25 (sd = 15.19). In the vocational training scenario the average reported SUS score was 78 (sd = 16.28) (see Figure 7). The SUS score of 75.25 out of possible 100 translate to a "good" and the SUS score of 78 out of possible 100 to a "good" or "excellent" usability according to Bangor et al. [2].

6.3.3 *Perceived Cognitive Load.*

The perceived cognitive loads were reported as a rTLX score of 20.4 (sd = 11.66) in the assistance scenario for cognitively impaired workers and a rTLX score of 28.7 (sd = 14.13) in the vocational training scenario for health-care professionals (see Figure 8).

6.3.4 *Qualitative Feedback.*

When asked if they would want to use the application for other tasks, in the assistance scenario for cognitively impaired workers, 6 participants answered "*yes*", stating that the glasses helped them understand what they had to do better (4 participants) and that they felt less pressured compared to the guidance of instructors (2 participants). The 4 participants that answered "*no*" stated as their reasons that they don't like technical things (2 participants) and that they perceived the glasses as uncomfortable because of their weight and the narrow field of view (2 participants).

In the vocational training scenario, 9 participants answered "*yes*". As their reasons they stated that the instructions gave them security and helped to make sure every step of a task was performed correctly and in correct order (6 participants), that the visual approach of the instructions was "memorable" (3 participants) and that they think learning a task using AR instructions would be faster (3 participants). The 1 participant that answered "*no*" stated as the reason the narrow field of view and that the glasses were occluding his vision.

In terms of additional qualitative feedback, in the assistance scenario for cognitively impaired workers, 6 participants stated that they perceived the glasses as heavy and that they would get headaches after prolonged use (2 participants), 3 participants reported that photo elements sometimes appeared blurry for them, and 3 participants stated that they had fun during the experiment.

In the vocational training scenario, 5 participants additionally noted that they liked the "detailed" explanation and 2 participants stated that they liked that the AR instructions can be used at their own pace as this takes away pressure. Furthermore, 3 participants reported that they sometimes had to search for the instructions if they were not immediately visible and 2 participants reported that the HoloLens was inconvenient to wear with their glasses. Additionally, 1 participant noted that the HoloLens became strenuous for the eyes after some use.

7 DISCUSSION

Through the combination of focus group interviews, the evaluation of the authoring tool itself and the evaluation of the authored instructions, several insights on the viability and usability of authorable AR instructions in work environments can be gained.

The instructors in the focus group interviews clearly stated that they perceive authorable AR as a viable option for assistance and training purposes in work environments. This positive perception was verified by the instructors using our proposed AR authoring tool to author a task in their scenario. Having actively evaluated the authoring, all instructors stated that they would like to use the tool in their scenario and that they perceive the tool as enabling enough to create concise and easy to understand AR instructions for their particular group of users. The usability score of 76, which translates to a "good" usability according to Bangor et al. [2], and the results of the user experience benchmark, where they perceived the tool as especially novel, stimulating and attractive, support their qualitative feedback and show that the provided tool is usable by them, even in its prototypical state. Their negative feedback was mainly aimed at the hardware of the HoloLens itself and regarding shortcomings that the already announced HoloLens 2 improves on (e.g. weight, field of view, usage while wearing glasses).

In line with the expectations stated by the instructors in the focus group interviews, workers were able to use the authored AR instructions for assistance and training without major problems or errors. They also rated the usability as "good" in the assistance scenario for cognitively impaired workers and between "good" and "excellent" in the health-care professional training scenario. They gave qualitative feedback indicating that they liked the visual approach, perceived the AR instructions as helpful and would like to use the AR instructions in their respective scenario for assistance or training purposes. The negative feedback was again mainly towards specific hardware limitations rather then the AR instructions itself. The average perceived cognitive load in form of rTLX score was 20.4 in the assistance and 28.7 in the training scenario. While task dependent and therefore not directly comparable, these results are well below average rTLX scores for medical (50.60) and mechanical (27.95) tasks identified in meta analyses [15] and therefore at least indicate that the authored AR instructions, even on the limiting hardware, did not overly increase cognitive load.

The students in the training scenario perceived the idea of authorable AR instructions as positive in the focus group interviews. However, they were somewhat reserved as to how fitting AR instructions would be for tasks they actually train. Therefore it is especially interesting that they ultimately rated the usability of the AR instructions the highest while also reporting the highest perceived cognitive load. This may at least partially be explained by the task chosen, which is more cognitively demanding and consequently the AR instructions might have been perceived as more helpful which finally resulted in a higher usability score.

7.1 Limitations & Future Work

The exploratory nature of this work inherently comes with limitations. While the primarily qualitative approach was chosen on purpose, as we believe it allowed for the most insights regarding the viability of authorable AR with the current state of the literature, further quantitative approaches are essential. Future work on authorable AR instructions should focus on long-term experiments, larger sample sizes and comparisons to preexisting conventional instructions in the real assistance and training scenarios they are evaluated in.

On the authoring tool itself, several improvements that are work in progress or were not evaluated in this work have to be completed. For example, the local storage should be improved towards cloud storage with user management functionality for personalised AR instructions as suggested by some instructors. The catalogue of 3D models should be expanded with the ability for third parties to add their own 3D models to the application. Also, in-view instructions should be added and evaluated as an alternative for the in-situ instructions. Additionally, the user should be given the choice between several different attention guiding techniques and additional types of AR elements (e.g. video or in-situ audio clips) could be considered.

Furthermore, the current work focused on the general applicability of the central part of an AR assistance or training system, the authoring and presentation of the step-wise instructions. Now that the approach has been demonstrated to be applicable, didactic concepts can be developed to integrate the technology into a coherent framework. For example, in real assistance and training scenarios it might not be desired to always display instructions but rather understand when and to what extend AR assistance is currently required [12, 39].

The authoring part should then also include wizards helping the instructor to identify the best way for instructing each individual step or a particular group of users with certain preferences. This, however, requires further and more systematic work on training and assistance performance of certain user groups, which in case of users with cognitive impairments is difficult to realize, as no two impairments are the same.

With upcoming technological advancements, such as the Microsoft HoloLens 2, which incorporates gesture recognition and eye-tracking capabilities, future work should explore the possibility to incorporate these for more intuitive interaction as previous research indicates eye-tracking [4] and gesture recognition [20] to be a promising choice for AR interaction. Additionally, the hardware improvements also enable action and state recognition and could therefore be used to detect problems or errors during the usage automatically [36]. A feedback mechanism, either implicit based on the performance or explicit by user interaction, could also inform the author to reconsider certain instructions, creating means for continuous improvements of the instruction sets.

8 CONCLUSION

An AR authoring tool combining the Microsoft HoloLens with a smartphone was developed, implemented and evaluated on a qualitative level based on related work, guidelines and feedback from focus groups. While the proposed tool was only a prototype, the tool itself and the created AR instructions were received well in terms of usability, user experience and qualitative feedback in both, the assistance scenario for cognitively impaired workers and the training scenario for health-care professionals.

The interaction between the smartphone and the HoloLens worked well for the instructors, showing up a viable alternative to the HoloLens-based interactions required with other approaches. In the near future, devices that pair non-autonomous smart glasses with smartphones are expected and the presented prototype demonstrates that this design choice might in case be advantageous, in particular for these information structuring tasks, which typically require interactions of high precision and the input of larger text segments.

Furthermore, the user feedback in our study revealed that both, instructors and workers, perceive authorable AR as a viable way to assist and train workers.

ACKNOWLEDGEMENTS

This research was supported by the Cluster of Excellence Cognitive Interaction Technology - "CITEC" (EXC 277) at Bielefeld University, which is funded by the German Research Foundation (DFG). The authors thank the *"Bildungszentrum Schopf (proWerk Bethel)"* and the *"SkillsLab (FH Bielefeld)"* for their cooperation to evaluate the application.

REFERENCES

[1] Maneesh Agrawala, Doantam Phan, Julie Heiser, John Haymaker, Jeff Klingner, Pat Hanrahan, and Barbara Tversky. 2003. Designing effective step-by-step assembly instructions. In *ACM Transactions on Graphics (TOG)*, Vol. 22. ACM, 828–837.

[2] Aaron Bangor, Philip Kortum, and James Miller. 2009. Determining what individual SUS scores mean: Adding an adjective rating scale. *Journal of usability studies* 4, 3 (2009), 114–123.

[3] Bruce Lawrence Berg, Howard Lune, and Howard Lune. 2004. *Qualitative research methods for the social sciences.* Vol. 5. Pearson Boston, MA.

[4] Jonas Blattgerste, Patrick Renner, and Thies Pfeiffer. 2018. Advantages of Eye-gaze over Head-gaze-based Selection in Virtual and Augmented Reality Under Varying Field of Views. In *Proceedings of the Workshop on Communication by Gaze Interaction (COGAIN '18).* ACM, New York, NY, USA, Article 1, 9 pages. https://doi.org/10.1145/3206343.3206349

[5] Jonas Blattgerste, Patrick Renner, and Thies Pfeiffer. 2019. Augmented Reality Action Assistance and Learning for Cognitively Impaired People. A Systematic Literature Review. In *The 12th PErvasive Technologies Related to Assistive Environments Conference (PETRAâĂŽ19).*

[6] Jonas Blattgerste, Patrick Renner, Benjamin Strenge, and Thies Pfeiffer. 2018. In-Situ Instructions Exceed Side-by-Side Instructions in Augmented Reality Assisted Assembly. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference (PETRA '18).* ACM, New York, NY, USA, 133–140. https://doi.org/10.1145/3197768.3197778

[7] Jonas Blattgerste, Benjamin Strenge, Patrick Renner, Thies Pfeiffer, and Kai Essig. 2017. Comparing Conventional and Augmented Reality Instructions for Manual Assembly Tasks. In *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '17).* ACM, New York, NY, USA, 75–82. https://doi.org/10.1145/3056540.3056547

[8] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.

[9] Philip Burnard. 1991. A method of analysing interview transcripts in qualitative research. *Nurse education today* 11, 6 (1991), 461–466.

[10] Sebastian Büttner, Markus Funk, Oliver Sand, and Carsten Röcker. 2016. Using head-mounted displays and in-situ projection for assistive systems: A comparison. In *Proceedings of the 9th ACM international conference on pervasive technologies related to assistive environments.* ACM, 44.

[11] Arindam Dey, Mark Billinghurst, Robert W Lindeman, and J Swan. 2018. A Systematic Review of 10 Years of Augmented Reality Usability Studies: 2005 to 2014. *Frontiers in Robotics and AI* 5 (2018), 37.

[12] Kai Essig, Benjamin Strenge, and Thomas Schack. 2016. ADAMAAS: towards smart glasses for mobile and personalized action assistance. In *Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments.* ACM, 46.

[13] Markus Funk, Thomas Kosch, Romina Kettner, Oliver Korn, and Albrecht Schmidt. 2016. motioneap: An overview of 4 years of combining industrial assembly with augmented reality for industry 4.0. In *Proceedings of the 16th international conference on knowledge technologies and datadriven business.*

[14] Markus Funk, Lars Lischke, Sven Mayer, Alireza Sahami Shirazi, and Albrecht Schmidt. 2018. Teach Me How! Interactive Assembly Instructions Using Demonstration and In-Situ Projection. In *Assistive Augmentation.* Springer, 49–73.

[15] Rebecca A Grier. 2015. How high is high? A meta-analysis of NASA-TLX global workload scores. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 59. SAGE Publications Sage CA: Los Angeles, CA, 1727–1731.

[16] Sinem Guven, Steven Feiner, and Ohan Oda. 2006. Mobile augmented reality interaction techniques for authoring situated media on-site. In *2006 IEEE/ACM International Symposium on Mixed and Augmented Reality.* IEEE, 235–236.

[17] Tracey E Hall, Charles A Hughes, and Melinda Filbert. 2000. Computer assisted instruction in reading for students with learning disabilities: A research synthesis. *Education and Treatment of Children* (2000), 173–193.

[18] Matthias Haringer and Holger T Regenbrecht. 2002. A pragmatic approach to augmented reality authoring. In *Proceedings. International Symposium on Mixed and Augmented Reality.* IEEE, 237–245.

[19] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.

[20] Wolfgang Hürst and Casper Van Wezel. 2013. Gesture-based interaction via finger tracking for mobile augmented reality. *Multimedia Tools and Applications* 62, 1 (2013), 233–258.

[21] Oliver Korn. 2012. Industrial playgrounds: how gamification helps to enrich work for elderly or impaired persons in production. In *Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems*. ACM, 313–316.

[22] Tobias Langlotz, Stefan Mooslechner, Stefanie Zollmann, Claus Degendorfer, Gerhard Reitmayr, and Dieter Schmalstieg. 2012. Sketching up the world: in situ authoring for mobile Augmented Reality. *Personal and ubiquitous computing* 16, 6 (2012), 623–630.

[23] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*. Springer, 63–76.

[24] Gun A Lee, Claudia Nelles, Mark Billinghurst, and Gerard Jounghyun Kim. 2004. Immersive authoring of tangible augmented reality applications. In *Proceedings of the 3rd IEEE/ACM international Symposium on Mixed and Augmented Reality*. IEEE Computer Society, 172–181.

[25] Kangdon Lee. 2012. Augmented reality in education and training. *TechTrends* 56, 2 (2012), 13–21.

[26] Bibeg Hang Limbu, Halszka Jarodzka, Roland Klemke, and Marcus Specht. 2018. Using sensors and augmented reality to train apprentices using recorded expert performance: A systematic literature review. *Educational Research Review* (2018).

[27] Blair MacIntyre, Maribeth Gandy, Steven Dow, and Jay David Bolter. 2004. DART: a toolkit for rapid design exploration of augmented reality experiences. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*. ACM, 197–206.

[28] David L Morgan. 1998. Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative health research* 8, 3 (1998), 362–376.

[29] Siti Soleha Muhammad Nizam, Rimaniza Zainal Abidin, Nurhazarifah Che Hashim, Meng Chun Lam, Haslina Arshad, and Nazatul Aini Abd Majid. 2018. A Review of Multimodal Interaction Technique in Augmented Reality Environment. *International Journal on Advanced Science, Engineering and Information Technology* 8, 4-2 (2018), 1460–1469.

[30] Riccardo Palmarini, John Ahmet Erkoyuncu, Rajkumar Roy, and Hosein Torabmostaedi. 2018. A systematic review of augmented reality applications in maintenance. *Robotics and Computer-Integrated Manufacturing* 49 (2018), 215–228.

[31] Wayne Piekarski and Bruce H. Thomas. 2002. The Tinmith System: Demonstrating New Techniques for Mobile Augmented Reality Modelling. *Aust. Comput. Sci. Commun.* 24, 4 (Jan. 2002), 61–70. http://dl.acm.org/citation.cfm?id=563997.563994

[32] Ivan Poupyrev, Desney S Tan, Mark Billinghurst, Hirokazu Kato, Holger Regenbrecht, and Nobuji Tetsutani. 2001. Tiles: A Mixed Reality Authoring Interface.. In *Interact*, Vol. 1. 334–341.

[33] Eduardo Quintana and Jesus Favela. 2013. Augmented reality annotations to assist persons with Alzheimers and their caregivers. *Personal and ubiquitous computing* 17, 6 (2013), 1105–1116.

[34] Patrick Renner and Thies Pfeiffer. 2017. Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 186–194.

[35] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2017. Construction of a Benchmark for the User Experience Questionnaire (UEQ). *IJIMAI* 4, 4 (2017), 40–44.

[36] Matthias Schröder and Helge Ritter. 2017. Deep Learning for Action Recognition in Augmented Reality Assistance Systems. In *ACM SIGGRAPH 2017 Posters (SIGGRAPH '17)*. ACM, New York, NY, USA, Article 75, 2 pages. https://doi.org/10.1145/3102163.3102191

[37] Hartmut Seichter, Julian Looser, and Mark Billinghurst. 2008. ComposAR: An intuitive tool for authoring AR applications. In *Proceedings of the 7th IEEE/ACM international symposium on mixed and augmented reality*. IEEE Computer Society, 177–178.

[38] Cledja Karina Rolim da SILVA. 2016. Design guidelines for generating augmented reality instructions. (2016).

[39] Benjamin Strenge, Ludwig Vogel, and Thomas Schack. 2018. Individualized cognitive assistance by smart glasses for manual assembly processes in industry. *Technische Unterstützungssysteme, die die Menschen wirklich wollen* (2018), 399.

[40] Thomas S Tullis and Jacqueline N Stetson. 2004. A comparison of questionnaires for assessing website usability. In *Usability professional association conference*, Vol. 1. Minneapolis, USA.

[41] Daniel W Turner III. 2010. Qualitative interview design: A practical guide for novice investigators. *The qualitative report* 15, 3 (2010), 754–760.

[42] Xin Wang, Soh K Ong, and Andrew YC Nee. 2016. A comprehensive survey of augmented reality assembly research. *Advances in Manufacturing* 4, 1 (2016), 1–22.

[43] Stefan Werrlich, Austino Daniel, Alexandra Ginger, Phuc-Anh Nguyen, and Gunther Notni. 2018. Comparing HMD-based and Paper-based Training. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 134–142.

[44] Stefan Werrlich, Phuc-Anh Nguyen, and Gunther Notni. 2018. Evaluating the training transfer of Head-Mounted Display based training for assembly tasks. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*. ACM, 297–302.

[45] Stefan Werrlich, Kai Nitsche, and Gunther Notni. 2017. Demand Analysis for an Augmented Reality Based Assembly Training. In *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '17)*. ACM, New York, NY, USA, 416–422. https://doi.org/10.1145/3056540.3076190

[46] Dennis Wolf, Daniel Besserer, Karolina Sejunaite, Matthias Riepe, and Enrico Rukzio. 2018. cARe: An Augmented Reality Support System for Dementia Patients. In *The 31st Annual ACM Symposium on User Interface Software and Technology Adjunct Proceedings*. ACM, 42–44.

[47] Dennis Wolf, John J Dudley, and Per Ola Kristensson. 2018. Performance Envelopes of In-Air Direct and Smartwatch Indirect Control for Head-Mounted Augmented Reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 347–354.

[48] Jürgen Zauner, Michael Haller, Alexander Brandl, and Werner Hartmann. 2003. Authoring of a Mixed Reality Assembly Instructor for Hierarchical Structures. In *Proceedings of the 2Nd IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '03)*. IEEE Computer Society, Washington, DC, USA, 237–. http://dl.acm.org/citation.cfm?id=946248.946783

[49] J Zhu, SK Ong, and AYC Nee. 2013. An authorable context-aware augmented reality system to assist the maintenance technicians. *The International Journal of Advanced Manufacturing Technology* 66, 9-12 (2013), 1699–1714.