8 Lenssembly: Authoring Assembly Instructions in Augmented Reality Using Programming-by-Demonstration

Thomas Kosch, Pascal Knierim, Mareike Kritzler, Daniel Beicht, and Florian Michahelles

Abstract

Managing the knowledge of assembly workers is crucial due to the valuable personal expertise of collected information over time that is hard to articulate. Unfortunately, the accumulated knowledge disappears when workers leave the company. Methods to record and transfer assembly knowledge between workers rarely exist due to the time-consuming documentation of assembly steps. This paper presents Lenssembly, a mobile augmented reality system utilizing programming-by-demonstration to record, detect, and generate assembly instruction sequences using a head-mounted display. The assembly instructions are automatically detected using a neural network, preventing the need for manual documentation and time-intensive content creation for each assembly step.

T. Kosch (✉)
Utrecht University, Utrecht, the Netherlands

HU Berlin, Berlin, Germany
e-mail: t.a.kosch@uu.nl

P. Knierim
Bundeswehr University Munich, Munich, Germany
e-mail: pascal.knierim@unibw.de

M. Kritzler
Siemens Corporate Technology, Berkeley, CA, USA
e-mail: mareike.kritzler@siemens.com

D. Beicht
LMU Munich, Munich, Germany

F. Michahelles
TU Wien, Vienna, Austria
e-mail: florian.michahelles@tuwien.ac.at

C. Röcker, S. Büttner (eds.), *Human-Technology Interaction*,
https://doi.org/10.1007/978-3-030-99235-4_8

In a user study (N = 12) with two different assembly tasks, participants favored the recording functionality of Lenssembly while conducting fewer errors and perceiving less task load than traditional paper instructions. We discuss the implications of our results and conclude how technologies create repositories for storing and transferring expert worker knowledge.

Keywords

Augmented reality · Artificial Intelligence · Programming-by-demonstration · Industry 4.0

8.1 Introduction

Assembly knowledge preservation has become a relevant factor in manual production lines. Lot sizes become smaller, which is attributed to a decrease in mass production and an increase in personalized production for individual customer needs [1]. Despite an increasing degree of automation, manual assembly workers are the driving force to maintain the trend of individual mass production [2]. Gone are the times where pure rote learning of instruction steps was sufficient to perform the manual assembly. Instead, individual production at assembly lines which are expected to rise [3]. Current practices require workers to memorize assembly instructions on-demand, requesting junior and senior workers to adapt their assembly procedures frequently to fulfill the unique product assembly requirements. The assembly knowledge is passed from a senior worker to a junior worker in verbal or written form. Common instruction modalities include verbal communication or printed instructions [4]. Such instruction modalities do not scale well with small sizes and frequently require senior workers to teach new instructions.

However, the senior workers' time is precious, and they may leave the company, effectively taking their accumulated expert knowledge with them. The current practice is to reverse engineer the assembly process since no documentation standards for assembly procedures exist. Here, junior workers are confronted with constant variations of assembly instructions that may lead to an increase in working memory demand and error rates [5–7]. The issues above are counterproductive towards an effective knowledge transfer between junior and senior workers.

Assistive systems at workplaces have been researched to ameliorate this effect [8]. Such assistive systems may use external displays [9], pick-by-light systems [10], or Augmented Reality (AR) systems, such as projections or Head-Mounted Displays (HMDs) [11] to support workers during their assembly. The feasibility of these systems has been the subject of various evaluations [12, 13] attesting to the positive effects of interactive worker assistance. However, one reason preventing assistive assembly systems from entering the production lines of enterprises is the extended complexity of recording and assembly instructions generation. Here, the initial configuration and the authoring of assembly instructions are time-consuming. Moreover, it requires senior workers proficient with the assembly procedure to prepare them alongside their working function in a company. Those requirements might result in an early decline of assistive assembly systems, although these systems offer promising solutions to convey assembly knowledge and optimize knowledge management.

This paper presents Lenssembly, an HMD-based prototype recording, generating, and displaying assembly instructions using a programming-by-demonstrating approach. Workers can record assembly instructions during assembly procedures through image data received from the integrated cameras. The assembly steps and associated objects are detected using a neural network translating the assembly instruction sequence into a spatial holographic representation. Lenssembly supports the modular recording of assembly steps, which can be stored to create an assembly instruction repository. Junior workers can then utilize this repository to view and assemble manufacturing procedures. We evaluate Lenssembly in a user study with two assembly tasks and 12 participants, comparing the usability of the teaching procedure and the assembly performance between Lenssembly and traditional paper instructions. We find that all participants could create assembly instructions and perform all assembly tasks successfully using Lenssembly. Furthermore, our results revealed that using Lenssembly contributes to a lower error rate and perceived task load compared to paper instructions. Finally, interviews revealed that Lenssembly is a suitable tool to preserve the assembly procedures of senior workers, which junior workers can use to learn new assembly procedures.

8.2 Contribution Statement

The contribution of our work is threefold: We (1) present Lenssembly, an AR-based assistive system that records and replays assembly instructions using a neural network through a user-driven programming-by-demonstration approach. We (2) conduct a user study ($N = 12$) with two different assembly tasks to evaluate the efficiency and usability of Lenssembly compared to traditional paper instructions. Finally, we (3) discuss how ubiquitous technologies benefit from our approach to populate knowledge repositories hosting assembly procedures for effective knowledge transfer between workers. We are confident that our work paves the way for pervasive knowledge documentation passed to other users utilizing ubiquitous technologies.

8.3 Related Work

In the following section, we first introduce related research regarding AR-supported assembly guidance, followed by an outline of assembly authoring and different object and action.

8.3.1 Augmented Reality Supported Assembly Guidance

The use of AR for assembly guidance instead of other assembly instruction modalities was already the subject of previous research. Boud et al. [14] presented the idea of using AR for manual assembly tasks. Specifically for the effectiveness of assembly task guidance, Boud et al. [14] compared five guidance methods, including conventional engineering drawings, immersive VR, and context-free AR. In a user study where participants had to conduct an assembly task using one of the different methods, they showed that VR and AR were outperforming the traditional 2D engineering drawings. The AR system was further rated as the most effective method. Interestingly, the assembly tasks were up to three times faster using AR compared to the VR methods and more than eight times faster than the traditional engineering drawings.

A series of studies showed the positive effects of augmented worker assistance. Henderson et al. [15], and Tang et al. [16] studied the advantages of AR compared to traditional knowledge and assembly instruction transfer methods, finding that AR-based guidance significantly reduced the task completion times, number of errors, and cognitive workload. Nilsson and Johansson [17] investigated the acceptance of AR instructions and confirmed the users' preference for AR supported instructions over traditional learning methods.

Researchers began to evaluate functional prototypes to evaluate augmented assembly instructions. A hand-held assembly system was presented by Billinghurst et al. [18]. A mobile phone is used as a see-through display to view complex models and detailed assembly instructions. Their study shows that AR with overlayed animations resulted in the lowest task completion time than static AR. Westerfield et al. [19] translated this concept to HMDs in a motherboard assembly scenario. Their research concludes that AR assembly guidance can significantly improve learning success. This hints towards a learning effect when lessons are frequently repeated in an interactive learning environment [20, 21].

Bannat et al. [22] presented how projection-based systems can support assembly workers. A camera and a projector mounted above the table displays the assembly instruction sequence. A camera is used to track the user's actions and verify the assembly. Inspired by the approach from Bannat et al., several projection-based systems were presented that improve the assembly performance at stationary and mobile workplaces for workers [12, 23–25] and workers with cognitive impairments [9, 26]. Furthermore, Kosch et al. [6, 27, 28] showed through the use of electroencephalography how the worker's cognitive workload is effectively reduced using in-situ AR. Hence, assistive assembly systems can be implicitly benchmarked in addition to already established assembly system performance metrics [29, 30]. Büttner et al. [31] investigated training duration and learning effects of AR-driven assembly support systems. Their findings conclude that participants show a faster assembly performance when using AR-based assistive technologies in the first 24 h. Their results imply that, while assistive technologies improve the assembly performance during initial training, there may be no improvements for long-term use.

8.3.2 Assembly Authoring, Object, and Action Recognition

The construction and order of assembly steps require a worker (i.e., the author) to demonstrate those steps while a system recognizes these actions. Marker-based approaches were initially used to accomplish object detection and tracking. Molineros et al. [32] addressed the sensing problem for object tracking and connection detection with the help of encoded markers. Each assembly object had a marker that was uniquely recognized and tracked. A previously computed assembly graph is used as a basis for the representation of all possible states of parts and feature descriptors. Computer vision algorithms are used to identify, track, and verify the assemblies and attachments in real-time using the marker. Gupta et al. [33] demonstrated an assembly authoring and guiding system for Lego Duplo bricks. The system uses color and depth information from a Kinect camera for object tracking and action inference. The functionality of the system is restricted to a table surface with four marked regions. The user can freely move the assembly object within the play area during the guidance and authoring task, where the object is tracked. A virtual replica is displayed on a monitor showing the assembly object in the same pose as the in-hand physical model. Before a new Lego Duplo brick is added to the physical model, it must first be placed in the "add box", where the size and color of the model are matched against a virtual model. If the attachment is not correctly recognized, the user can restart the detection process by placing his hand in a recheck box. Finally, objects can be removed from the physical model by placing the detached object in the remove box. For the alignment of the digital replica with the physical model, a transformation between the camera point cloud and the virtual point cloud is calculated with an iterative closest point algorithm. New attachments and detachments are detected by calculating a belief distribution value over several frames, describing addition and removal candidates, since single poses (i.e., single frames) contain insufficient information. This approach can be combined with recent studies that utilize eye tracking to understand how assembly instruction information is accessed by the individual user [34].

Neural networks have become a popular method to estimate object and their states for further processing in AR. Su et al. [35] proposed a method to estimate the pose of objects for AR applications. They use a neural network to detect the pose and state of objects for further use in AR applications. Roitberg et al. [36] presented a vision-based approach for recognizing human activities in industrial environments using deep learning. The system enables the recognition of hand activities like selecting points and objects grasping and coarser activities such as assembly and processing of workpieces. Multimodality is used to create a single output by combining multiple sensors. For this, data from a Kinect v2 for human skeleton tracking, an Asus Xtion Pro for object recognition, and a Leap Motion for detailed hand skeleton tracking are fused with a self-developed control framework for data synchronizing. Activities are classified in multiple abstraction levels using hierarchical hidden Markov models after feature selection and dimensionality reduction. Agrawala et al. [37] proposed design principles on how to display assembly instruction sequences for users. Consequently, Büttner et al. presented a system that records assembly instructions of Lego Duplo bricks using an RGB and depth sensor [38, 39]. The prototype records the user's actions and visualizes the assembly instructions to the user. However, the assembly instruction reconstruction is limited to a stationary setting and relies on complementary object colors to enable a robust recognition. While the added depth sensor provides a more robust detection based on the objects' 3D constitution, it may not work reliably for similar objects with similar colors.

Neural networks have gained an increase in popularity for object and action detection recently. Shinde et al. [40] present an approach where human actions are recognized using the YOLO object detection framework. Therefore, they trained the object detector with images consisting of appropriate labeled actions. When an action label is detected with a given threshold over multiple frames, the associated action is classified for this period. An object detection algorithm for action classification allowed the classification of activities with a small number of images. It showed that even single images could be sufficient to recognize the action. The work of Bhattacharya et al. [41] introduced an approach for generating AR work instructions by expert demonstration. The execution of the procedure is first recorded and then processed to create the learning environment. For this purpose, the system consists of two parts: the demonstration and refinement phases. In the demonstration phase, a static near-range 3D sensor captures a predefined surface. New steps are detected by continuously checking whether a hand is in the image. The current point cloud is compared with a previous one and checked for a new object when the hand disappears from the image. Subsequently, an algorithm tries to identify the movement of the newly detected object to create an accurate animation. Finally, the refinement phase allows us to modify the recorded steps and add additional text, images, and videos. However, this approach is limited to stationary settings as a near-range sensor must be placed in a fixed position. Recently, Kong et al. [42] presented TutorialLens, an AR authoring and demonstration system to create instructions for the operation of user interfaces. Users can author interface tutorials using a computer vision-based demonstration recording the 3D coordinates of finger positions. TutorialLens reproduces these finger movements for users who are not proficient with the user interface through a see-through device. Users are then hinted at the correct interaction via AR.

Previous research showed how interactive worker assistance could be used to increase assembly efficiency and productivity. However, at the same time, assembly instructions for those assistive systems need to be created, which is a laborious task. Here, past research looked into activity and action recognition to ease the authoring of the assembly instructions. However, previous work required manual interaction with an assistive system to teach assembly instruction procedures or was functional under controlled conditions. We close this gap by presenting Lenssembly. This mobile assembly instruction authoring system ubiquitously records the worker's interaction with workpieces to automatically

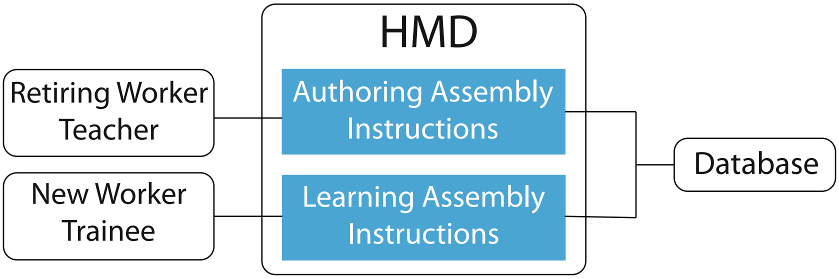

Fig. 8.1 Application structure of Lenssembly. An Authoring Mode enables to create and author new assembly instruction procedures. The learning procedure uses the recorded instructions to transfer this knowledge to new workers via AR. A Playback Mode displays the assembly instructions in AR and verifies if the assembly is conducted correctly

generate assembly instruction sequences, where objects are detected using a neural network.

8.4 Lenssembly: An Assembly Authoring and Playback System

Lenssembly runs as a mobile setup on a Microsoft HoloLens as HMD. Here, Lenssembly features an Authoring Mode and Playback Mode to capture assembly steps of a senior worker and replay them to trainees or junior workers. Figure 8.1 depicts the concept of Lenssembly. Both, the Authoring Mode and Playback Mode use the front-facing camera of an HMD (i.e., a HoloLens) to capture changes on the workplace to learn the order of new assembly steps (i.e., in the Authoring Mode) or show a digital representation of the next assembly step (i.e., in the Playback Mode). Objects and the worker's hands are detected using a trained neural network. We use the YOLOv3 [43] algorithm to detect and track objects. YOLOv3 applies a single neural network to the full image, making it suitable for the efficient real-time detection of objects. Objects must be annotated beforehand in pictures fed into the neural network, resulting in a model embedded into applications. The following section describe the details of the Authoring Mode and Playback Mode.

8.4.1 Authoring Mode: Expert Authoring and Recording Systems

The Authoring Mode allows expert and senior workers to record the order of assembly steps using an automatically "programming-by-demonstration" approach. The recording system allows workers to author new assembly task workflows and store them persistently in a data model by demonstrating the assembly steps in a workplace. This includes assembly steps and hands, tools, and actions that require attachments between several components. The Authoring Mode system utilizes a neural network to detect the sequence

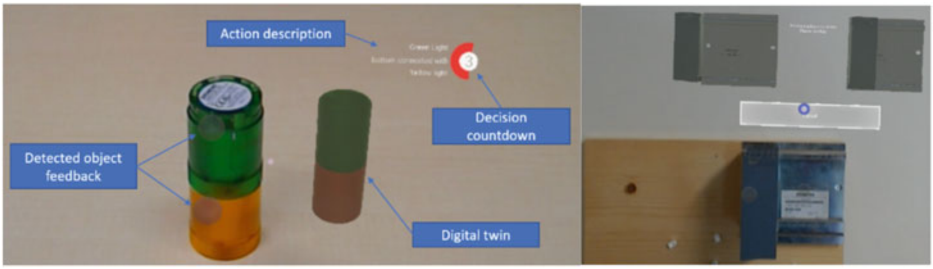

Fig. 8.2 Procedure step suggestion process for a detected connected object in the Authoring Mode. **Left:** A connection between the yellow and green object was detected and is added to the assembly procedure after the decision timer expires. **Right:** Conflict resolution menu models two detected actions for the same objects. The user must choose one of the two actions by clicking on the corresponding model

of the assembly steps and worker actions. The user can manually proceed or wait until a decision countdown expires, upon which the current assembly instruction step is saved. The worker receives continuous feedback about the learned assembly procedure and displayed digital representation. The previously prepared digital twin is saved into the app and displayed upon detecting the associated real-world object. All objects are captured using the point-of-view camera of the HoloLens. Figure 8.2 depicts an illustration of the Authoring Mode view.

The worker can choose an existing assembly task or create a new one upon starting Lenssembly. A virtual keyboard or voice commands are used throughout the whole training procedure to provide text input. The worker can create a new work procedure for each step and perform the respective assembly. Lenssembly provides constant feedback on how the final representation will look like for the trainee. The process enables experts to see what the system detected and to intervene if necessary. First, the procedure step is always described as a triplet in the upper right corner of the HMD. In the case of an attachment action, a digital twin is created next to the associated physical objects of the action. To create a digital twin, we use pre-generated 3D models of the individual captured objects utilizing the CAD models to create a 3D representation in AR. The 3D representations of the digital twins are prepared beforehand and displayed when the associated objects' constellations are detected. A five-second timer is started upon detecting an object. The assembly procedure is saved when the timer expires, and subsequently, the next step can be performed—the timer restarts when corrections to the assembly are made. The current assembly step training procedure is canceled when the front-facing camera loses the tracking of the objects. If multiple actions are detected during the same assembly step, the expert worker can review and correct the assembly steps using a conflict resolution menu (see Fig. 8.2).

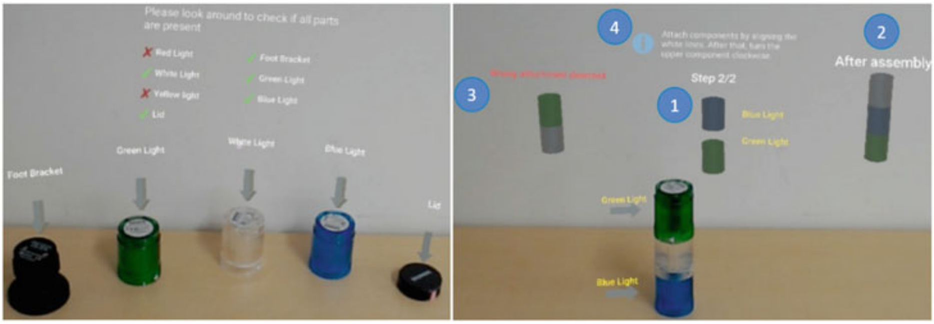

Fig. 8.3 Lenssembly checks the available parts prior the assembly. **Left:** A checklist is displayed and an arrow with a description denotes the part. The assembly begins when all parts are available. **Right:** Lenssembly displays (1) the current procedure step, (2) the pre-trained result, (3) potential wrong actions, and (4) textual information messages that support the assembly

8.4.2 Playback Mode: Trainee Replay and Learning System

Previously trained assembly instructions can be replayed using a Playback Mode functionality. Lenssembly displays all authored assembly procedures for selection to the user. A feedback location (i.e., the worktable) has to be selected by the user to begin with the assembly process using a drag and drop gesture. Then, a part list appears that visually checks if the assembly parts are available in the workplace (see Fig. 8.3). The assembly starts afterward. A previously prepared 3D rendering of a CAD model is rendered next to the physical assembly object. A description above the object provides additional details about how to perform the assembly. Generated animations are played to display how attachments must be performed (e.g., through a predefined blinking arrow displayed above the detected object). Also, arrows depict which objects have to be attached. The assembled objects turn green, and the application continues with the next assembly step when a correct assembly is detected. Wrong actions are displayed on the left side of the workplace to inform the user about potential corrections of their assembly. Again, all objects are tracked using the integrated point-of-view camera of the HoloLens. Figure 8.3) displays how the Playback Mode is displayed to the user.

8.5 Evaluation of Lenssembly Through a User Study

We conduct an evaluation of the Authoring Mode and Playback Mode system of Lenssembly. We explain the used assembly tasks, provide details about the trained model to detect workpieces and describe the methodology.

Fig. 8.4 The signaling column (**left**) and PLC task (**right**) that were used to evaluate Lenssembly. Both assembly tasks require a screwdriver

8.5.1 Assembly Tasks

We selected two assembly scenarios to evaluate Lenssembly in a user-centric study. First, we use YoloV3 to train a neural network that detects the assembly components, tools, and worker's hands. Second, we utilize the assembly of a signaling column and programmable logic controller for the user study.

Signaling Column Assembly The signaling column assembly task consists of seven components that are stacked on top of each other (see Fig. 8.4). At the lower end is a bracket, and the upper back is closed by a lid. Every component, except the screwdriver, has one or two attachment points located on the top or bottom of the object where other fitting parts can be attached. The bracket can be screwed on to fasten the signaling column. This assembly task was chosen because of the various possible results and procedures that can lead to the same goal. Actions, such as the connections between the components, the worker's hand, and the screwdriver, need to be detected by Lenssembly. We have selected the signaling column assembly task since it can be varied through the order of the attachments without changing the assembly task itself.

Programmable Logic Controller Assembly The programmable logic controller (PLC) task contains five hardware components (see Fig. 8.4) that require the worker to mount several components on a mounting rail. The components are placed next to each other. The PLC task is more complex than the signaling column task since the modules placed next to each other and on the components cannot be exchanged. In addition, it requires the frequent use of the screwdriver since all modules need to be fastened.

8.5.2 Data Set Collection and Model Training

We recorded videos of all assembly workpieces in the first step to acquiring image data of the objects. Videos were made from different angles and directions to retrieve a diverse set of images. Furthermore, all videos were taken against diverse backgrounds and under other lighting conditions. We recorded the videos with 30 frames per second. We extracted every fifth frame of the videos for the labeling process to get a good distribution of multi-angled images. We manually removed blurred images and replaced them with adjacent frames from the video. We ensured that all objects in the data set were represented approximately equally often during the data collection. In total, we extracted over 5600 frames from the recorded videos. Each captured frame contained one or multiple objects. The data set included 17 classes from the two previously introduced assembly tasks. A list of all classes with their associated assembly task can be found in Table 8.1. Tools (i.e., a screwdriver) and hands are listed extra since they are part of both tasks.

Data annotation and labeling is the process of labeling data for supervised learning machine learning methods. This involves the object localization inside a frame and their respective manual labeling. We manually labeled the objects with bounding boxes in the image using the open-source tool OpenLabeling. OpenLabeling generates the annotation and associated labels in the appropriate YOLOv3 textfile format. Darknet2 is used as an open-source implementation for YOLOv3. Each file stores information about all bounding boxes of the corresponding image, describing their respective class id and position (i.e., xCenter, yCenter, width, and height), and uses normalized pixel coordinates between 0 and 1. Finally, we achieved a good distribution with only the Hand class having significantly more pictures~ (see Table 8.1). After generating the basic truth data, the annotated data set was split into a train and a test data set. We divided the data into 85–15 splits using a random-based approach. YOLOv3 resized the input images to 608x608 pixels. We used the previous labeled images for training, resulting in 11,467 training instances. The number of filters was set to 66 in the YOLOv3 configuration file. The final model was trained for 176k iterations until a training loss of less than 0.1 was reached, resulting in a loss of 0.06.

8.5.3 Methodology

We evaluate Lenssembly using a mixed-method study design opting for a within-between-subject design. We divided the participants into two groups. The first group (N = 6) was asked to assemble the signaling column and PLC unit using the Playback Mode of Lenssembly. Afterward, the participants filled a NASA-TLX questionnaire and custom Likert scales accompanied by a semi-structured interview. The second group (N = 6) was asked to assemble the signaling column and PLC unit using paper instructions. The participants were asked to fill a NASA-TLX questionnaire and custom Likert scales similar to the first group. In addition to the assembly, the second group was asked to build the

Table 8.1 Final label distribution of the workpieces, tools, and hands

Class	Instances	Assembly task
Yellow light	659	Signaling column
Red light	684	Signaling column
Green light	693	Signaling column
White light	663	Signaling column
Blue light	663	Signaling column
Lid	686	Signaling column
Bracket	674	Signaling column
Load current supply (lid closed)	689	PLC
CPU (lid closed)	660	PLC
Digital input/output module (lid closed)	655	PLC
Mounting rail	580	PLC
Load current supply (lid opened)	641	PLC
CPU (lid opened)	650	PLC
Digital input/output module (lid opened)	665	PLC
U-connector	632	PLC
Screwdriver	667	Tool
Hand	906	Tool
Total	11,467	

signaling column using the Playback Mode to gain additional qualitative feedback about the mode. Finally, we conducted a semi-structured interview. We counterbalanced the assembly conditions (i.e., assembly of signaling column and PLC unit) according to the balanced Latin square. After the assembly, all participants of both groups were invited to train and test a new assembling procedure of a signaling column using the Authoring Mode to gain additional qualitative feedback.

We compare the assembly efficiency of Lenssembly using paper instructions during the participant's assembly using the Playback Mode. Printed paper instructions are a standardized modality to convey assembly instructions that have been used in research to compare novel assembly instruction modalities [44]. The paper instructions included one instruction step per page. The page has to be flipped to continue with the following assembly instruction. Pictures were generated using available CAD models of the corresponding workpieces. Also, both the recorded assembly procedures in AR and the paper instructions contained written text to support the assembly. We used the Microsoft HoloLens as HMD. Figure 8.5) the study workplace. We outline the independent and dependent variables in the following.

Independent Variables We employ the assembly instruction modalities as a factor with two levels (i.e., AR with Lenssembly and paper instructions). Participants were either assembling the signaling column task or PLC task using Lenssembly or using paper

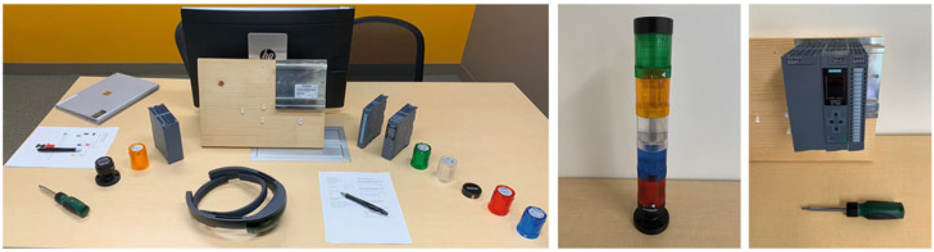

Fig. 8.5 **Left:** Study setup with the signaling column and PLC assembly task. Printed paper instructions were used as baseline. The finally assembled **(middle)** signaling column and **(right)** PLC unit

instructions depending on the group. In addition, all participants were teaching the signaling column using the Authoring Mode of Lenssembly after their assembly.

Dependent Variables We measure the task completion time for every single assembly step during assembly with the paper instructions and Playback Mode of Lenssembly. We measured the number of errors made for each assembly task. An error is always counted when the assembly step itself or the mounting of parts was conducted erroneously (e.g., wrong order of attachments). Here, we subdivide the errors between independent and dependent errors. We specify independent errors as wrong attachment steps or missing assembly objects. For example, independent errors are subsequently performed without the user's dependent error being recognized by the user. Each participant was asked to fill in a NASA Task Load Index (NASA-TLX) [45] questionnaire after each assembly procedure. Afterward, we conducted semi-structured interviews to gain additional feedback about the Playback Mode and Authoring Mode.

8.5.4 Procedure

We greeted the participants and provided them with a written description of the study to ground each participant regarding their intention. The participants were familiarized with the study setup after they provided their informed consent. Furthermore, our participants provided their demographic data, previous AR experience, and knowledge about manual assembly. A 1-min video an introduction of the Playback Mode of Lenssembly. It demonstrated the basic UI elements, controls, and the execution of hand tool actions. The study started afterward with the assembly using either Lenssembly or paper instructions according to the counterbalanced order.

In the first part of the study, our participants started with assembling the first workpiece after putting on the HoloLens or viewing the printed paper instruction. Here, the participant was either asked to assemble the workpiece (see Fig. 8.5 for the fully assembled result). For the signaling column, the participants were attaching the subparts of the signaling column

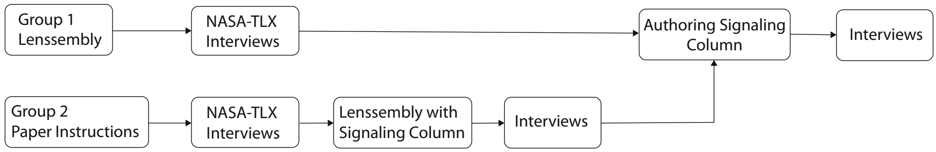

Fig. 8.6 Illustrated study procedure. We use mixed within-between-subject design where both groups experience the assembly using either Lenssembly or paper instructions. All participants author a signaling column after the assembly using the Authoring Mode. Interviews were held between the sessions

until it resulted in the whole signaling column that consisted of seven objects and six assembly steps. On the other hand, the PLC task consists of five objects and fifteen procedure steps that frequently involve tool actions. In addition, participants filled a NASA-TLX questionnaire after each assembly. Finally, all participants were authoring a signaling column using the Authoring Mode.

The second part of the study focused on the Authoring Mode feature of Lenssembly. First, the participants were asked to train an assembly procedure of the signaling column. Again, an introductory video explained how the Authoring Mode works, how it suggests detected actions, and how they can intervene to correct the system if necessary. Next, the participants were invited to build the signaling column in their fashion due to the modular components. The participants were asked to test their training procedures afterward. Finally, a semi-structured interview was conducted that examined the usability and user acceptance of the Authoring Mode. Figure 8.6 depicts the study procedure.

8.5.5 Participants

We recruited 12 volunteers (three female, nine male) to participate in the user study via mailing lists. The participants' age ranged between 18 and 55 years. Four participants had no previous experience with AR, where six participants reported rare experience with AR. Two participants were using AR frequently. Six participants had prior experience with the HoloLens. None of the participants had previous experience with manual assembly.

8.6 Results

In the following, we report the results of the task completion time, the number of errors, and subjectively perceived task load.

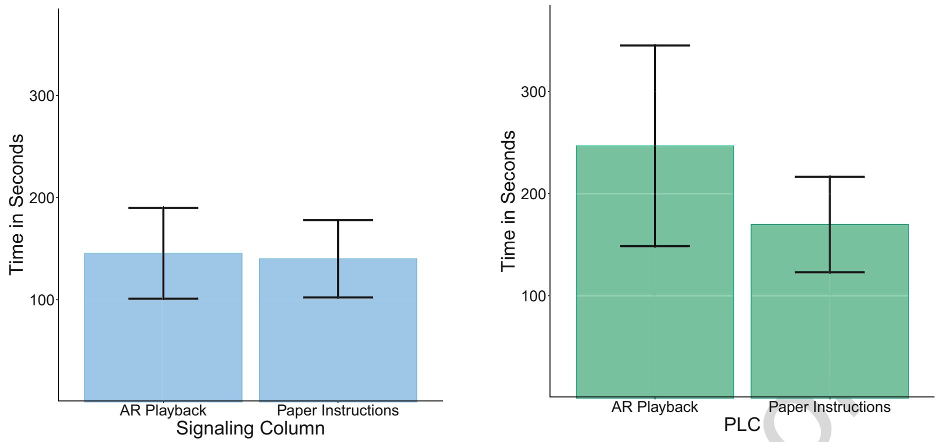

Fig. 8.7 Task completion times for the signaling column **(left)** and the PLC assembly (**right**). Paper instructions require less time for complex assembly procedures. The error bars depict the standard error

8.6.1 Task Completion Time

Paper instructions required less time compared to instructions displayed in the Playback Mode when assembling the signaling column (paper instructions: M = 140.00s, SD = 37.80s, AR-based instructions: M = 145.55s, SD = 44.49s). For the PLC task, participants require less time using paper instructions (M = 170.00s, SD = 46.84s) compared to the Playback Mode (M = 247.00s, SD = 98.33s). Potential reasons for longer task completion times using the AR-based instructions can be that participants were unused to the HMD itself and the potential waiting times between the assembly steps to display the next instruction. Figure 8.7 illustrates the mean task completion times.

8.6.2 Number of Errors and Task Load

We counted the number of errors conducted with paper and the employed AR-based instructions during the Playback Mode. We describe the number of independent and dependent errors. The documentation method for the signaling column task had 0.83 errors and 1.00 independent errors on average, resulting in 1.83 errors. No error was recorded from participants using the Playback Mode. For the PLC assembly task, we measured an averaged independent error rate of 0.33, with no independent error occurring during Playback Mode (overall error rate: M = 0.33). However, the documentation manual had an independent error rate of 1.83 and a dependent error rate of 0.33 (overall error rate: M = 2.16). Figure 8.8 shows the averaged number of errors. Furthermore, we analyze the NASA-TLX questionnaire to quantify the participants' subjectively perceived task load.

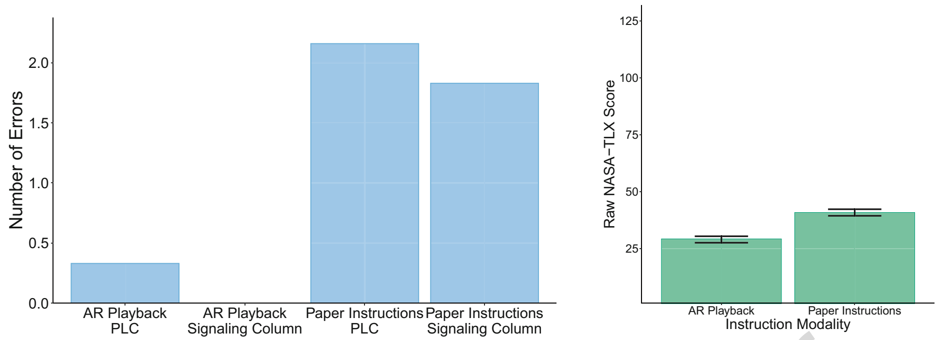

Fig. 8.8 **Left:** Averaged number of errors for both instruction modalities. Paper instructions elicit a larger number of paper instructions compared to the Playback Mode. **Right:** Raw NASA-TLX scores for both instruction modalities. The Playback Mode requires less workload compared to paper instructions

We find a higher level of workload for paper instructions ($M = 40.84$, $SD = 1.44$) compared to instructions displayed in the Playback Mode ($M = 29$, $SD = 1.42$). Figure 8.8 shows the TLX scores between the Playback Mode and paper instructions.

8.6.3 Qualitative Results

The following section examines the findings of the semi-structured interviews. Most participants liked that they had their hands free during the assembly and did not have to occupy their hands with physical instructions. However, one participant (P9) stated that it is hard to do the actual assembly when handling the AR application and the assembly task. The participant referred to the tool actions, where the search for tools "[. . .] leads to abnormal head movements." (P9). This is mainly attributed to the narrow field of view of the HoloLens. However, we see this as a technical limitation that will be resolved in the future. Asking about potential helpful features that are present in the application, eight participants stated that they liked the arrow feedback (P1, P2, P5–P7, P10–P12). Two participants added that it helped them to identify the objects even if you do not know what they look like (P7, P12). This was useful in the PLC assembly task. However, one participant found there was sometimes "[. . .] too much information with the arrows" (P3) which led to confusion. P7 liked that you can look at the rendered 3D models from all perspectives and that you can even look into them, providing an experience which would not be possible with documentation or videos. The participants also liked the fact that the application tells you when a task is completed. One participant stated that "It makes you feel good. You can tick off the task, continue, and know that the previous step was correct." (P10). Some participants were disturbed that certain information was often not in the field of view of the assembly environment. We asked what worked well and what caused problems during the learning process. Four participants stated that they had problems carrying out the screwing action because their hands were not recognized (P2, P3, P9, P10). Two participants explained that they often did not see the visual feedback when the system successfully recognized an object. Thus, they missed the next step in the process (P3, P9), resulting in higher task completion times. One participant stated that this is probably the case because the confirmation text on the HMD lies on a different level than the physical objects (P3). The participant also stated that "[. . .] if you concentrate on the assembly and your eyes are focused on it, you simply overlook information on the HMD." (P3). We asked if they liked that the system checks for the correctly assembled objects before proceeding to the next step. The participants liked the concept since it detects and catches errors early on. One participant thought it is essential for work where a certain level of safety must be maintained (P11). We continued to ask technical questions about the HMD. First, we asked what features of the HoloLens would need to be improved to increase the application's usability. Ten participants stated that the narrow field of view is one of the main limitations (P1–P8, P11, P12). Five participants had problems with the resolution as well as the contrast and found it hard to distinguish between certain colors.

Finally, we were asking questions about the perceived utility and usefulness of the Authoring Mode. Most of the participants perceived the recording process as intuitive. One participant stated that "[...] this is the next step in learning assembly tasks after documentations and video instructions." (P3). All participants stated that they either wanted to set the time for the recording countdown manually or use a dedicated gesture, button, or voice command. One participant would like to disassemble objects later in the recording to undo previously recorded procedure steps (P9). Overall, all participants could imagine using an HMD application similar to the developed one in the future to store their assembly knowledge persistently.

8.7 Discussion

Can we use ubiquitous technologies to store and preserve assembly knowledge? The presented user study results show that our participants could learn the recorded assembly tasks and train new assembly procedures using our system. This section discusses the results, addresses the remaining challenges, and lays out the vision of user-generated knowledge repositories.

8.7.1 Lenssembly Requires More Time than Paper Instructions

While the signaling column assembly task did not differ in completion times between the AR application and the documentation, the PLC task showed a descriptive difference. However, paper instructions outperformed AR-based instructions in terms of task completion times during the PLC task. One reason was the tool actions, where the tool has to be registered by the HMD, hence taking more time than using the tool right away. Furthermore, some users intuitively grabbed the screwdriver without looking at the tool to perform the pickup action, ignoring the previous demonstration of handling these actions that led to the need to register the tool again. This is consistent with the collected qualitative statements, where participants stated thfat the special performance of tool actions could lead to unusual head movements. Also, some participants had problems performing hand tool actions since their hands were not reliably detected. One reason could be that the data set contained hand images from people with a similar handshape. A more diverse set of hand images can lead to a more robust model. The additional actions that should be performed with the tools were partially forgotten and led to problems during execution, contributing to longer execution times.

Although the overall assembly with AR took longer than paper instructions, participants were faster going through the checklist for the available parts. The participants stated that the arrows and named labels at the location of physical objects are more convenient than searching for objects using images in the paper instructions. Verifying the availability of the required assembly parts was favored by the participants. However, another issue that arose during the assembly was the narrow field of view of the camera, preventing a successful detection of the performed assembly. We believe this can be circumvented by adding additional wearable cameras around the HMD or improvements to the existing camera. Participants also complained that the field of view is limited. We see this as a technical limitation that is likely to be resolved with future releases of HMDs (e.g., the HoloLens 2).

8.7.2 Lenssembly Elicits Fewer Errors and Less Task Load

We observed that text written in the paper instructions was more likely to be ignored than visually represented in AR. We attribute the higher number of errors for paper instructions compared to AR-based instructions to these observations. The participants confirmed this, who either did not read textual representations of assembly instructions or missed them. Our results imply that if time is not a critical factor, AR-based instructions lead to fewer errors, hence maintaining a quality level of the assembled product compared to paper-based instructions. Contrary to this, paper-based instructions can be used if time is a critical component. An analogy can be drawn toward the perceived task load. It is advisable to use AR-based instructions when the user's perceived workload should be kept minimal. For example, safety-critical assembly scenarios benefit from the Playback Mode used by junior workers.

Finally, we believe Lenssembly is successful in conveying assembly instructions to junior workers. Overall, the results show that the learning environment with automatic object detection and feedback generation led to successful assembly procedures for all participants. It reduces the error rate compared to printed documentation by providing feedback and preventing dependent errors. The current restrictions lead to the same or slower completion times compared to documentation. However, most participants stated that learning with Lenssembly was more engaging and that after the assembly, instruction was known. The HMD is not needed anymore after completing several assembly trials. We believe that Lenssembly provides an entry into the assembly of complex workpieces, which is not needed anymore after the worker gets used to the assembly procedures.

8.7.3 Recording Assembly Instructions

The evaluation of the recording subsystem showed that the participants were able to record a custom signaling column assembly procedure. All participants were able to record their assembly procedures after a short introduction. Similarities between colors, such as red and yellow, were challenging to distinguish for the deployed classifier. However, since the error only occurred in one room, the problem could have been different lighting conditions. Most participants built the signaling column with all available components and wanted to see their automatically generated learning environment afterward.

8.7.4 Limitations

The presented implementation and the study design are prone to certain limitations. The neural network can only detect an object that has been trained before. This requires laborious labeling of data which was performed manually during this study. Furthermore, this could have caused some noise in the training data since the bounding boxes do not cover the exact shape of the assembly workpieces. However, 3D representations of workpieces are usually available as a CAD model. We believe that these models can be exploited to generate a neural network for the parts and the respective assembly order, hence obviating the need for manual labeling. Another limitation affecting the study is the limited field of view of the HoloLens itself. However, this is a technical limitation that will be improved through future developments. A final fundamental limitation is the manual preparation and association of the digital twins towards the detected objects. The arrangement of the different hologram parts has to be stored into Lenssembly directly to be displayed upon detection of the assembled object. However, we are confident that future work can use the detected object model to generate the digital twin on the fly. Finally, we acknowledge that our sample partly represents persons who have experience with the HoloLens. This experience can skew our results in favor of the HoloLens. However, we are confident workers will achieve a similar performance after a brief settling-in period.

8.7.5 Future Work

Our research provides the first step of creating a knowledge repository that is not limited to the interactive representation of assembly instructions alone. In the first step, we want to utilize existing 3D CAD models to automatically generate a neural network that can detect single workpieces in an assembly procedure to create assembly instructions automatically. Furthermore, we envision creating a user-generated repository that conveys assembly instructions enriched by individual assembly styles that optimize previously known assembly procedures. Also, we will investigate potential learning effects that emerge during the use of AR-based instructions. Here, we are interested in differences in using assembly instruction in different modalities, including the required length of use and assembly performance. Finally, we will incorporate the suggestions made by the participants with a subsequent evaluation of learning effects for various assembly scenarios.

8.8 Conclusion

This paper presents Lenssembly, an AR-based system that enables the knowledge transfer of assembly instructions between senior and junior workers. We developed an AR application for an HMD that enables the flexible authoring of assembly instructions. Workers can record their assembly procedure using a programming-by-demonstration approach, reducing the complexity of recording assembly instructions while generating instructions simultaneously. As a result, the laborious content creation of assembly instructions is no longer required. In a user study with 12 participants that evaluates the Authoring Mode and Playback Mode, we find a reduced number of assembly errors and self-reported task load when using Lenssembly compared to paper instructions. However, Lenssembly elicited a higher task completion time compared to paper instructions. Interviews revealed that training new assembly procedures was a pleasant experience that all participants have successfully conducted. We conclude that Lenssembly eases the authoring of further assembly instructions for assistive systems and supports junior workers to learn new assembly instruction procedures. Our work describes how recent advances in artificial intelligence can be used to preserve assembly knowledge by creating user-generated assembly sequences that others can retrieve for learning purposes. We are confident that our work paves the way for future usable assembly systems that automatically generate assembly instructions.

References

1. Wang, Y., Ma, H. S., Yang, J. H., & Wang, K. S. (2017). Industry 4.0: A way from mass customization to mass personalization production. *Advances in Manufacturing, 5*(4), 311320. https://doi.org/10.1007/s40436-017-0204-7

2. Fellmann, M., Robert, S., Büttner, S., Mucha, H., & Röcker, C. (2017). Towards a framework for assistance systems to support work processes in smart factories. In A. Holzinger, P. Kieseberg, A. M. Tjoa, & E. Weippl (Eds.), *Machine learning and knowledge extraction* (p. 5968). Springer.
3. Yin, Y., Stecke, K. E., & Li, D. (2018). The evolution of production systems from industry 2.0 through industry 4.0. *International Journal of Production Research, 56*(1–2), 848861. https://doi.org/10.1080/00207543.2017.1403664
4. Gjeldum, N., Salah, B., Aljinovic, A., & Khan, S. (2020). Utilization of Industry 4.0 related equipment in assembly line balancing procedure. *Processes, 8*(7), 864. https://doi.org/10.3390/pr8070864
5. Brettel, M., Friederichsen, N., Keller, M., & Rosenberg, M. (2014). How virtualization, decentralization and network building change the manufacturing landscape: An industry 4.0 perspective. *International Journal of Mechanical, Industrial Science and Engineering, 8*(1), 3744.
6. Kosch, T., Funk, M., Schmidt, A., & Chuang, L. L. (2018). Identifying cognitive assistance with mobile electroencephalography: A case study with in-situ projections for manual assembly. *Proceedings of the ACM on Human Computer Interaction, 2*, 11. https://doi.org/10.1145/3229093
7. Rüßmann, M., Lorenz, M., Gerbert, P., Waldner, M., Justus, J., Engel, P., & Harnisch, M. (2015). Industry 4.0: The future of productivity and growth in manufacturing industries. *Boston Consulting Group, 9*(1), 5489.
8. Büttner, S., Mucha, H., Funk, M., Kosch, T., Aehnelt, M., Robert, S., & Röcker, C. (2017). The design space of augmented and virtual reality applications for assistive environments in manufacturing: A visual approach. In *Proceedings of the 10th ACM international conference on PErvasive technologies related to assistive environments*. ACM. https://doi.org/10.1145/3056540.3076193
9. Funk, M., Mayer, S., & Schmidt, A. (2015). Using in-situ projection to support cognitively impaired workers at the workplace. In *Proceedings of the 17th international ACM SIGACCESS conference on computers & accessibility*. ACM. https://doi.org/10.1145/2700648.2809853
10. Guo, A., Raghu, S., Xie, X., Ismail, S., Luo, X., Simoneau, J., Gilliland, S., Baumann, H., Southern, C., & Starner, T. (2014). A comparison of order picking assisted by head-up display (HUD), cart-mounted display (CMD), light, and paper pick list. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers, ISWC'14* (p. 7178). Association for Computing Machinery. https://doi.org/10.1145/2634317.2634321
11. Büttner, S., Funk, M., Sand, O., & Röcker, C. (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, PETRA'16*. Association for Computing Machinery. https://doi.org/10.1145/2910674.2910679
12. Funk, M., Baechler, A., Baechler, L., Korn, O., Krieger, C., Heidenreich, T., & Schmidt, A. (2015). Comparing projected in-situ feedback at the manual assembly workplace with impaired workers. In *Proceedings of the 8th international conference on PErvasive technologies related to assistive environments*. ACM. https://doi.org/10.1145/2769493.2769496
13. Zheng, X. S., Foucault, C., Matos da Silva, P., Dasari, S., Yang, T., & Goose, S. (2015). Eye-wearable technology for machine maintenance: Effects of display position and hands-free operation. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems, CHI'15* (p. 21252134). Association for Computing Machinery. https://doi.org/10.1145/2702123.2702305
14. Boud, A. C., Haniff, D. J., Baber, C., & Steiner, S. (1999). Virtual reality and augmented reality as a training tool for assembly tasks. In *1999 IEEE International Conference on Information Visualization (Cat. No. PR00210)* (p. 3236). IEEE.

15. Henderson, S. J., & Feiner, S. K. (2011). Augmented reality in the psychomotor phase of a procedural task. In *2011 10th IEEE International Symposium on Mixed and Augmented Reality* (p. 191200). IEEE.
16. Tang, A., Owen, C., Biocca, F., & Mou, W. (2003). Comparative effectiveness of augmented reality in object assembly. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (p. 7380). ACM.
17. Nilsson, S., & Johansson, B. (2008). Acceptance of augmented reality instructions in a real work setting. In *CHI'08 extended abstracts on Human factors in computing systems* (p. 20252032). ACM.
18. Billinghurst, M., Hakkarainen, M., & Woodward, C. (2008). Augmented assembly using a mobile phone. In *Proceedings of the 7th international conference on mobile title suppressed due to excessive length 21 and ubiquitous multimedia* (p. 8487). Association for Computing Machinery. https://doi.org/10.1145/1543137.1543153
19. Westerfield, G., Mitrovic, A., & Billinghurst, M. (2015). Intelligent augmented reality training for motherboard assembly. *International Journal of Artificial Intelligence in Education, 25*(1), 157172. https://doi.org/10.1007/s40593-014-0032-x
20. Lindgren, R., & Johnson-Glenberg, M. (2013). Emboldened by embodiment: Six precepts for research on embodied learning and mixed reality. *Educational Researcher, 42*(8), 445452. https://doi.org/10.3102/0013189X13511661
21. Werrlich, S., Daniel, A., Ginger, A., Nguyen, P., & Notni, G. (2018). Comparing HMD-based and paper-based training. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (p. 134142). ACM. https://doi.org/10.1109/ISMAR.2018.00046
22. Bannat, A., Wallhoff, F., Rigoll, G., Friesdorf, F., Bubb, H., Stork, S., Müller, H. J., Schubö, A., Wiesbeck, M., Zäh, M. F., et al. (2008). Towards optimal worker assistance: A framework for adaptive selection and presentation of assembly instructions. In *Proceedings of the 1st international workshop on cognition for technical systems*. Cotesys.
23. Funk, M., Bächler, A., Bächler, L., Kosch, T., Heidenreich, T., & Schmidt, A. (2017). Working with augmented reality? A long-term analysis of in-situ instructions at the assembly workplace. In *Proceedings of the 10th ACM international conference on PErvasive technologies related to assistive environments*. ACM. https://doi.org/10.1145/3056540.3056548
24. Funk, M., Kosch, T., Kettner, R., Korn, O., & Schmidt, A. (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* (p. 4). ACM.
25. Funk, M., Sahami Shirazi, A., Mayer, S., Lischke, L., & Schmidt, A. (2015). Pick from here! - An interactive mobile cart using in-situ projection for order picking. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*. ACM. https://doi.org/10.1145/2750858.2804268
26. Kosch, T., Kettner, R., Funk, M., & Schmidt, A. (2016). Comparing tactile, auditory, and visual assembly error-feedback for workers with cognitive impairments. In *Proceedings of the 18th international ACM SIGACCESS conference on computers & accessibility*. ACM. https://doi.org/10.1145/2982142.2982157
27. Kosch, T., & Chuang, L. (2018). Investigating the impact of assistive technologies on working memory load in manual assembly through electroencephalography. In *2nd international neuroergonomics conference: The brain at work and in everyday life*. Frontiers Research Foundation.
28. Kosch, T., Karolus, J., Ha, H., & Schmidt, A. (2019). Your skin resists: Exploring electrodermal activity as workload indicator during manual assembly. In *Proceedings of the ACM SIGCHI*

Symposium on Engineering Interactive Computing Systems, EICS'19. ACM. https://doi.org/10.1145/3319499.3328230
29. Funk, M., Kosch, T., Greenwald, S. W., & Schmidt, A. (2015). A benchmark for interactive augmented reality instructions for assembly tasks. In *Proceedings of the 14th international conference on mobile and ubiquitous multimedia*. ACM. https://doi.org/10.1145/2836041.2836067
30. Kosch, T., Abdelrahman, Y., Funk, M., & Schmidt, A. (2017). One size does not fit all - Challenges of providing interactive worker assistance in industrial settings. In *Proceedings of the 2017 ACM international joint conference on pervasive and ubiquitous computing*. ACM. https://doi.org/10.1145/3123024.3124395
31. Büttner, S., Prilla, M., & Röcker, C. (2020). Augmented reality training for industrial assembly work - Are projection-based AR assistive systems an appropriate tool for assembly training? In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI'20* (p. 112). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376720
32. Molineros, J. M. (2001). *Computer vision and augmented reality for guiding assembly*. Springer.
33. Gupta, A., Fox, D., Curless, B., & Cohen, M. (2012). Duplotrack: A real-time system for authoring and guiding Duplo block assembly. In *Proceedings of the 25th annual ACM symposium on user interface software and technology* (p. 389402). ACM.
34. Heinz, M., Büttner, S., & Röcker, C. (2020). Exploring users' eye movements when using projection-based assembly assistive systems. In N. Streitz & S. Konomi (Eds.), *Distributed, ambient and pervasive interactions* (p. 259272). Springer.
35. Su, Y., Rambach, J., Minaskan, N., Lesur, P., Pagani, A., & Stricker, D. (2019). Deep multistate object pose estimation for augmented reality assembly. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)* (p. 222227). ACM. https://doi.org/10.1109/ISMAR-Adjunct.2019.00-42
36. Roitberg, A., Somani, N., Perzylo, A., Rickert, M., & Knoll, A. (2015). Multimodal human activity recognition for industrial manufacturing processes in robotic workcells. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (p. 259266). ACM.
37. Agrawala, M., Phan, D., Heiser, J., Haymaker, J., Klingner, J., Hanrahan, P., & Tversky, B. (2003). Designing effective step-by-step assembly instructions. *ACM Transactions on Graphics, 22*(3), 828837. https://doi.org/10.1145/882262.882352
38. Büttner, S., Peda, A., Heinz, M., & Röcker, C. (2020). Teaching by demonstrating how smart assistive systems can learn from users. In N. Streitz & S. Konomi (Eds.), *Distributed, ambient and pervasive interactions* (p. 153163). Springer.
39. Funk, M., Lischke, L., Mayer, S., Shirazi, A. S., & Schmidt, A. (2018). *Teach me how! Interactive assembly instructions using demonstration and in-situ projection* (p. 4973). Springer. https://doi.org/10.1007/978-981-10-6404-34
40. Shinde, S., Kothari, A., & Gupta, V. (2018). Yolo based human action recognition and localization. *Procedia Computer Science, 133*, 831838.
41. Bhattacharya, B., & Winer, E. H. (2019). Augmented reality via expert demonstration authoring (AREDA). *Computers in Industry, 105*, 6179.
42. Kong, J., Sabha, D., Bigham, J. P., Pavel, A., & Guo, A. (2021). Tutoriallens: Authoring interactive augmented reality tutorials through narration and demonstration. In *Title suppressed due to excessive length 23 Symposium on Spatial User Interaction. SUI'21*. Association for Computing Machinery. https://doi.org/10.1145/3485279.3485289
43. Redmon, J., & Farhadi, A. (2018). *Yolov3: An incremental improvement*. ACM.
44. Blattgerste, J., Strenge, B., Renner, P., Pfeiffer, T., & Essig, K. (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 (p. 7582). Association for Computing Machinery. https://doi.org/10.1145/3056540.3056547

45. Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (task load index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload, advances in psychology* (Vol. 52, pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9