Time is money! Evaluating Augmented Reality Instructions for Time-Critical Assembly Tasks

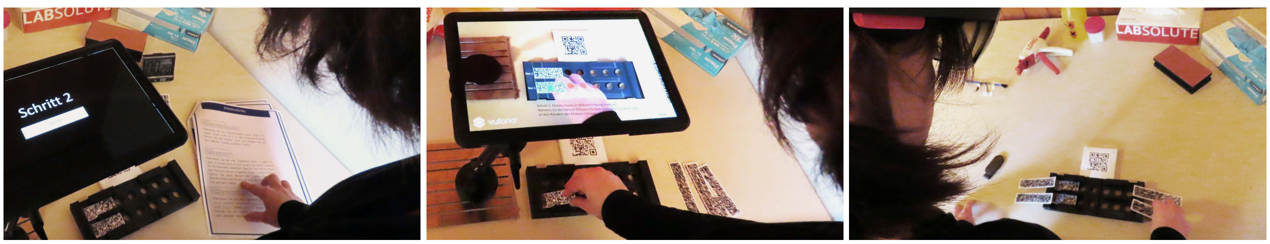
Jannike Illing
OFFIS Institute for Information Technology
Smart Human Robot Collaboration
Oldenburg, Germany
jannike.illing@offis.de

Philipp Klinke
OFFIS Institute for Information Technology
Safety & Security Oriented Analysis
Oldenburg, Germany
philipp.klinke@offis.de

Uwe Grünefeld
University of Duisburg-Essen
Essen, Germany
uwe.gruenefeld@uni-due.de

Max Pfingsthorn
OFFIS Institute for Information Technology
Smart Human Robot Collaboration
Oldenburg, Germany
max.pfingsthorn@offis.de

Wilko Heuten
OFFIS Institute for Information Technology
Interactive Systems
Oldenburg, Germany
wilko.heuten@offis.de



(a) Paper instructions. **(b) Tablet instructions.** **(c) Head-mounted instructions.**

Figure 1: Instructions evaluated on paper, tablet (Samsung Galaxy Tab S4), and head-mounted device (Microsoft HoloLens).

ABSTRACT

Manual assembly tasks require workers to precisely assemble parts in 3D space. Often additional time pressure increases the complexity of these tasks even further (e.g., adhesive bonding processes). Therefore, we investigate how Augmented Reality (AR) can improve workers' performance in time and spatial dependent process steps. In a user study, we compare three conditions: instructions presented on (a) paper, (b) a camera-based see-through tablet, and (c) a head-mounted AR device. For instructions we used selected work steps from a standardized adhesive bonding process as a representative for common time-critical assembly tasks. We found that instructions in AR can improve the performance and understanding of time and spatial factors. The tablet instruction condition showed the best subjective results among the participants, which can increase motivation, particularly among less-experienced workers.

MUM 2020, November 22–25, 2020, Essen, Germany

ACM ISBN 978-1-4503-8870-2/20/11…$15.00
https://doi.org/10.1145/3428361.3428398

CCS CONCEPTS

• **Applied computing** → **Computer-aided manufacturing**; • **Human-centered computing** → *Mixed / augmented reality*; *User studies*.

KEYWORDS

augmented reality, user study, computer-aided manufacturing, manual processes, worker support

ACM Reference Format:
Jannike Illing, Philipp Klinke, Uwe Grünefeld, Max Pfingsthorn, and Wilko Heuten. 2020. Time is money! Evaluating Augmented Reality Instructions for Time-Critical Assembly Tasks. In *19th International Conference on Mobile and Ubiquitous Multimedia (MUM 2020), November 22–25, 2020, Essen, Germany.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3428361.3428398

1 INTRODUCTION

Manual assembly tasks are often characterized by aligning and joining parts in three-dimensional space. In many cases, workers additionally have to consider given time constraints while putting these parts together. Therefore, the performance and quality of such assembly tasks depend on both, precise spatial positioning and acting in specific time frames. A representative example of such a task is the adhesive bonding[1] process in which spatial aspects and

[1] Adhesive bonding (also known as glue bonding or gluing) describes the process of joining two surfaces together.

time constraints play an essential role. For example, spacers must be placed in the correct position, and the joining parts placed precisely on top of each other. In addition, depending on the adhesive joints, there are different flash-off, pot life, and setting times, which must be strictly adhered to during the process [9]. Otherwise, the adhesive joint will fail, rendering the assembly unsuccessful.

Adhesive bonding processes are characterized by high safety requirements such as the observance of special hygiene conditions and wearing protective clothing. Nowadays, especially in small and medium-sized enterprises (SMEs), bonding processes are characterized by many manual work steps, which are presented to the worker on paper-based instruction sheets. Manually adjustable timers are used to measure time and duration for the execution of tasks and waiting times. Multiple processes are often conducted simultaneously, so it is challenging for the worker to keep an overview of the different complicated steps and execution times.

In manual assembly and manufacturing, Augmented Reality (AR) has proven to be a helpful technology because it enables more pleasant and efficient working processes in terms of simulation, support, and guidance [17, 20]. Furthermore, less paperwork is needed because of the dynamic real-time information provided [3].

AR devices' fundamental idea is to change users' perceived reality by superimposing digital information onto the real environment [2, 26]. This superimposed information allows real-world objects to be annotated by AR, which is useful for providing users with additional information [18]. In particular, AR allows the accurate and precise mapping of real objects at future positions with a virtual model. For example, parts of an assembly process can be virtually positioned in advance, and a user gets an exact (spatial) impression of the required assembly position before the actual assembly process starts [1, 11, 16]. Hence, we believe that AR is a promising technology for providing users with instructions for time-critical assembly tasks. However, to our knowledge, it is not clear which AR device works best to provide users with time-critical assembly instructions.

In this paper, we compare *paper*, *tablet*, and *head-mounted* instructions in a laboratory user study. Both the *tablet* and *head-mounted* instructions are displayed in AR, while the *paper* instructions represent the established analog approach and serves as a baseline condition. Our results contribute to the support of workers, dealing with time-critical processes, in which exact spatial positioning/assembly of parts plays an important role. In addition, high safety/hygiene requirements must be observed in adhesive bonding processes, which make precise work even more difficult. To sum up, we want to investigate the possibilities of support by AR and reduce possible mistakes such as the non-observance of different times and the wrong spatial positioning of spacers and joining parts. The specific contribution of this paper is an evaluation of AR instructions in a laboratory user study compared to a paper-based baseline condition.

2 RELATED WORK

In the following, we discuss the related work regarding: (1) AR form factors and (2) AR work instructions. We aim to identify the advantages of already proven AR form factors and AR instructions. Nevertheless, to the best of our knowledge, no previous work has explored time-critical assistance in AR.

2.1 Comparing AR Form Factors

In previous work, different AR form factors have been explored ranging from smartphones [1, 21] and tablets [8, 11, 16, 22] to head-mounted [8, 11, 16, 21, 22] and projected AR devices [1, 8].

If the tasks required users to perform an assembly task, they preferred projected AR over AR with smartphones [1], tablets [8] and head-mounted displays [8]. In these tasks, it was found that task execution with projected AR was significantly faster [1, 8] and the error rate was lower compared to head-mounted [8] or smartphone [1] instructions. In addition, higher cognitive stress was reported for the head-mounted [8] instructions. The tablet or paper instructions by Funk et al. [8] hindered the participants with both hands.

It is also interesting to compare head-mounted displays and tablets when performing assembly tasks [11] or in use in industrial shop floors [22] without the use of projected AR. At Hoover [11], user feedback favored the tablet application, whereby the head-mounted display used resulted in significantly faster assembly times. At Syberfeldt et al. [22], in turn, users responded positively to the optical see-through head-mounted display due to the hands-free operation. However, in both studies, users frequently reported that the head-mounted display felt heavy after some usage time [11, 22]. It should also be noted that the head-mounted display used by Syberfeldt et al. [22] is not up to current standards. A possible explanation for the preference of the tablet at Hoover [11] is provided by Serubugo et al. [21], who compared wearable (smart glasses) and handheld (smartphone) devices as a platform for an AR museum guide. It was found that the operation of handheld devices is more intuitive and familiar compared to wearable devices.

Another comparative study, but without industrial context, was conducted by Plasson et al. [16] to evaluate the performance and usability of tablet-based and see-through head-mounted display-based interaction techniques. Here, the head-mounted display was perceived by the participants as faster and physically less strenuous, with the direct touch technique less affected by small targets and occlusion.

2.2 Work instructions in Augmented Reality

Several studies have been published on the topic of AR, quantifying the advantages of AR instructions to traditional instructions such as paper-based instructions [4, 12, 23, 24], video instructions [14], digital 2D instructions [23], or expert tutorials [24]. Furthermore, Sanna et al. [19] reported that handheld AR devices contribute to error and time minimization as well as to a positive user experience.

A first concept for using AR to display work instructions was proposed by Caudell and Mizell [6] already in 1992. They developed an early prototype of AR and postulated that a transparent and trackable head-mounted display could be used to display work instructions, thus reducing the paper load.

The usage of AR instructions in the execution of assembly tasks showed a lower mental workload [12] and a reduction of the error rate [4, 12] compared to traditional isometric drawings [12] or conventional pictorial instructions [4]. In turn, in execution time, Hou et al. [12] and Blattgerste et al. [4] report different results. For Hou et al. [12], the assembly time was reduced by half using the AR instructions, while for Blattgerste et al. [4], the fastest way to complete the task was to use the paper instructions.

Besides paper instructions, Wiedenmaier et al. [24] added a tutorial by an expert for comparison. Thereby, support of assembly processes by AR with tasks of varying degrees of difficulty was examined. Their results show that the assembly times varied. The AR support proved to be more suitable for difficult tasks than the paper manual, while the use of a paper manual for simpler tasks did not differ significantly from the AR support. Tasks performed under the guidance of an expert were completed most quickly.

A comparison between video and AR assistance in manual assembly was performed by Loch et al. [14]. Here the AR assistance showed a significantly lower number of errors and a better evaluation in terms of time and mental stress.

Another interesting result is provided by Tang et al. [23], who compared AR with digital 2D instructions and paper instructions. A lower mental stress was reported and a lower error rate was reported when using the AR solution. The results confirm the statements of Hou et al. [12] and Loch et al. [14] regarding a lower mental stress and Hou et al. [12], Blattgerste et al. [4] and Loch et al. [14] regarding reduction of the error rate when using AR.

In addition to comparative studies, Pathomaree and Charoenseang [15] enhanced the skill transfer in the assembly task with a training system in AR. They used virtual objects and graphical instructions for advising the user with the assembly steps and the targeted positions in assembly task. Their results show that a training system in AR increase the transferability and transfer effectiveness ratio. In addition, their system reduced assembly completion times and the number of assembly steps. The users were very satisfied with this kind of system.

3 CONTEXT OF USE ANALYSIS

In order to analyze the context of an adhesive bonding process, an observation study was first carried out in a real training environment (see Figure 2) in accordance with the Human-Centered Design Process (HCD). The aim was to identify the tools, protective clothing, materials, and equipment used. The laboratory premises were divided into four areas: seminar area, work area, safety area, and storage area.

Thereby, we could observe at which steps of an adhesive bonding process possible errors and uncertainties on the part of the trainees occur. The trainees came from different companies that have integrated or want to integrate "bonding" as a sub-process. Another important observation was the different times that the trainees have to observe during the adhesive bonding process. In addition, various parameters that can influence the times. For example, the pot life is dependent on both the given (room) temperature and the amount to be applied.

The observation was followed by interviews with bonding technology experts, who were able to give us a real picture of the production and assembly operations and identify frequent sources of error. With the help of the interviews, we were also able to determine the need for technical support.

From the observation and the expert interviews, we could determine that the logging and the process steps are often incomplete. This can be both a lack of knowledge on the part of the workers and carelessness in execution. These causes can be, for example, the incorrect spatial positioning of the spacer/joining parts. In addition, mistakes are often made in pre-treatment, such as cleaning or primer application. Some errors, such as not meeting specific time requirements, can lead to the adhesive bond to fail eventually. In addition to process-related errors, cognitive stress has also been identified. Workers must have a high level of knowledge about safety regulations, chemicals, process specifications, and times, and often in series production, they must be able to carry out all process steps entirely and precisely.

4 DESIGN

Following the context of use analysis, our analysis of previous research and practice began to determine relevant preliminary results and extract various human-technology interaction concepts for the execution of assembly tasks.

For the visualization of paper-based (adhesive) instructions and the possibility of hands-free interaction, the head-mounted AR, especially with the advanced Microsoft HoloLens [13] device, has proven to be the most suitable. To enable hands-free interaction with the tablet, it was attached to a fixed mount with a ball joint. Thus the tablet could "drive" over the (gluing) device[2] (similar to a magnifying glass). In this way, it was possible to prevent holding the tablet from hindering the execution of the task with both hands, as is the case with paper instructions [8, 25].

4.1 Assembly instructions

Our selected process steps were derived both from the observations of the context analysis (see section 3) and from the requirements for adhesive bonding processes defined in DIN 2304-1:2019-10 [7]. Thus, a standard-compliant representation could be implemented, which contains all necessary information for the execution of the individual work steps. For our study, we selected four work steps from an overall process that are particularly prone to errors (see Figure 3. The steps are used in training courses and are therefore particularly suitable, as our study was carried out by participants without any experience in adhesive technology.

Step 1 (preparation): Two spacers must be positioned at the correct position of the (gluing) device. For this purpose, two positions[3] are displayed on the left side of the (gluing) device in the AR view.

Step 2 (minimum timer): Four joining parts have to be flashed off and positioned in the correct position. After positioning, one timer per joining part starts. A joining part must not be processed before the timer has expired. In the AR view, four possible positions are displayed. As soon as a part is placed on one of these positions, a timer for this part starts.

Step 3 (preparation): Every joining part must be positioned laterally on the (gluing) device for the next working step to be directly

[2]A gluing device describes a custom-made fixture in which spacer and joining parts can be inserted with an exact fit.

[3]All positions are represented in the AR view by holograms of the corresponding parts.

Figure 2: Context of use analysis of the training environment.

accessible. In the AR view, a total of 4 positions are displayed (2 positions each on the left and right side of the (gluing) device). The joining parts must be placed on these positions.

Step 4 (maximum timer): A joining part must be glued[4] and placed in the correct position in the (gluing) device. One timer starts for each adhered part. Another (non-glued) part must be placed on the (glued) part in the correct position before the timer expires. In the AR view, two positions are shown on the right side of the (gluing) device. When a part is placed on one of the positions, a timer starts and the position of the counterpart is displayed on the left in the (gluing) device. As soon as the counterpart is placed on this position, the timer is stopped. Example: If the first joining part is placed on the front right, the second joining part must be placed on the front left within the displayed time.

5 EXPERIMENT

To investigate the possibilities of support by AR, we conducted a user study. In this study, we implemented selected steps of an adhesive bonding process with AR in the form of a *tablet* and *head-mounted* variation and traditional *paper* instructions and then compared the performance of the participants.

5.1 Study Design

The experiment was conducted as a mixed design with two groups to exclude possible learning effects between the *AR* instructions. Both groups received *paper* instructions as the baseline condition. We investigated the differences between the *paper* and AR (*tablet* or *head-mounted*) instructions within one group. Between the two groups, we were able to test how the two *AR* instructions differ from each other. A participant was assigned to one of these groups. For example, the third participant was assigned to the head-mounted group and first performed the *paper* instructions and then the *head-mounted* instructions. Each group following a complete counterbalanced design (see Table 1).

Table 1: Mixed design with counter-balancing.

Group	Participant	Instructions
Tablet group	1, 5, 9, 13, 17	Paper + Tablet
	2, 6, 10, 14, 18	Tablet + Paper
Head-mounted group	3, 7, 11, 15, 19	Paper + Head-mounted
	4, 8, 12, 16, 20	Head-mounted + Paper

We used quantitative methods to evaluate performance, taking total process time, timer errors, positions of spacers and joining parts, task performance, and subjective measures as our dependent variables. To evaluate the performance of the users by using the different variants of instructions, we conducted an independent controlled laboratory study with traditional *paper* instructions and AR (*tablet* and *head-mounted*) instructions.

For this study we asked: **(RQ) To what extent can the performance of the participants in terms of working time, workload and error rate be improved by using AR for time and spacial dependent process steps?**

H_1 We expect that tablet-based and head-mounted AR will increase the understanding and execution of the individual process steps compared to paper instructions because holograms directly indicate the placement of the joining parts

[4]For safety reasons, the part of the actual gluing is only fictitious.

Figure 3: Work steps of the assembly instructions in Augmented Reality.

at the correct position of the (gluing) device and timers no longer have to be started manually.

H_2 We expect that the tablet solution will be preferred by the user over the head-mounted solution due to the assembly task at a fixed workplace because no additional adjustment to the human being (putting on something) is required.

5.2 Apparatus

We set up an empty office room with open windows and switched on lights to ensure that the AR devices could identify the spacer and joining parts at all times. Our *AR* instructions are displayed on the Microsoft HoloLens and Samsung Galaxy Tab S4.

For the implementation we used Unity(v.2018.4.5)[5], a 3D game development platform. Vuforia (v.8.3.8)[6] marker detection is used to detect the (gluing) device, the spacer, and joining parts.

Distance measurement. To measure the distance we placed a reference point rf (displayed as green cubes on each spacer) (see Figure 4 or Figure 3). Then we placed the spacer in the correct position into the (gluing) device and placed another (optimal) reference point[7] orf (displayed as red cubes) on the (gluing) device in the exact same position where rf was located (after placing it in the device). The distance between rf and orf could then determined with the help of the euclidean distance. When the cubes were almost on top of each other, the value of the distance ranged around 0 which would also mean, that the spacer was placed close to the optimal position.

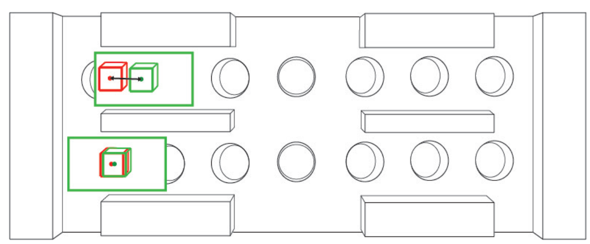


Figure 4: Procedure for measuring the distance.

[5]https://unity.com, last retrieved March 31, 2019

[6]https://developer.vuforia.com, last retrieved March 27, 2019

[7]The optimal reference points were invisible to the participants.

5.3 Procedure

The experiment took place in an empty office room, with the participants standing or sitting in front of the (gluing) device. Each experiment started with a detailed explanation of the experiment procedure so that the participants could try out the conditions, the working area, and the task description in advance. Each condition consisted of four (working) steps of an adhesive bonding process, which had to be performed one after the other (see Figure 3). At the end of the last instruction, a short questionnaire eliciting demographics, experience with instructions, experience with AR, and subjective measures were administered. Each participant needed about 25 minutes to complete the experiment.

Paper instructions. In addition to the *paper* instructions, the participants were given four preset laboratory timers in order to correctly perform steps 2 (minimum timer) and 4 (maximum timer). The timers had to be started manually. In addition, a tablet was placed over the (gluing) device to record the experiment through log files. A black screen with the respective step and a button was displayed on the tablet. The decision whether a step was finished was made by the participants themselves. A participant tapped the button on the tablet after completing a step.

Tablet instructions. The participants received a tablet representing the working instructions by AR. The tablet was in a fixed holder with a ball joint - so the tablet could easily be "moved" over the working area. The required positioning of the spacer or joining parts was represented by holograms. The decision whether a step was finished was made by the participants themselves. To do this, a participant tapped the button on the tablet after completing a step. The minimum and maximum time started automatically as soon as the part was correctly positioned.

Head-mounted instructions. The visualization and procedure did not differ from the *tablet* instructions. Instead of the tablet, participants received a Microsoft HoloLens that also displayed the work instructions in AR. However, to confirm a completed step, a clicker was used instead of a button.

5.4 Participants

We recruited a total of 20 volunteer participants[8] through public and online advertisements.

Tablet group: 10 participants (6 male, 4 female), aged between 21 and 62 (M=34.70, SD=14.61). Nobody suffered from color vision impairment, five had corrected-to-normal vision and five had normal vision.

Head-mounted group: 10 participants (6 male, 4 female), aged between 19 and 46 (M=27.70, SD=10.13). One participant suffered from color vision impairment, six had corrected-to-normal vision and four had normal vision.

We asked the participants to rate their experience with AR on a 5-point-Likert-scale. We also asked to rate their experience with manual work instructions on a 5-point-Likert-scale. In both groups, the participants indicated that they had limited to medium experience with AR (Md=2.00, IQR=1.00) and that they had medium to very high experience with work instructions (Md=3.50, IQR=1.00).

5.5 Results

Spacer Distance. To find out how the participants understand the different instructions, we compared the distance from the optimal position of the spacers in the (gluing) device with the (real) position. This allowed us to determine how the participants placed the spacers in the (gluing) device (see section 5.2). Therefore, we recorded all positions of the spacers (including position and rotation) during the study. The median distance per condition for the tablet group in ascending order are: *tablet* instructions (Md=0.05 cm, IQR=0.05 cm) and *paper* instructions (Md=0.06 cm, IQR=0.11 cm). The median distance per condition for the head-mounted group in ascending order are: *paper* instructions (Md=0.14 cm, IQR=0.32 cm) and *head-mounted* instructions (Md=0.14 cm, IQR=0.06 cm). The distances are compared in Figure 5.

Figure 5: The median spacer distance per group without outliers.

A Shapiro-Wilk-Test showed that our data are not normally distributed ($p < 0.001$). We used a Wilcoxon Signed-rank test for not normally distributed data and non-parametric, related groups (within subjects). This showed no significant differences between the distances in the tablet group (W_{min}=87.0, z=-0.672, p=0.502, r=0.150[9]) as in the head-mounted group (W_{min}=74.0, z=-1.157, p=0.247, r=0.259). For non-parametric, unrelated groups (between subjects) with not normally distributed data, we used a Mann-Withney U-test. This showed a significant difference between the distances of the *tablet* and *head-mounted* instructions (U_{min}=82.0, z=3.178, p=0.001, r=0.711) and no significant difference between the distances of the *paper* instructions of both groups (U_{min}=169.0, z=0.825, p=0.409, r=0.184).

Joining Parts Distance. We compared the distance from the joining parts in the (gluing) device. The median distance per condition for the tablet group in ascending order are: *paper* instructions (Md=0.04 cm, IQR=0.07 cm) and *tablet* instructions (Md=0.05 cm,

[8]For mean effect sizes of (d=0.60), at least 74 observations are necessary, which requires testing at least 10 participants (for each condition we have 4 trials per participant). We calculated this value with G*Power under Wilcoxon Mann-Whitney U-test for unmatched pairs (α=0.05 and 1-β=0.80).

[9](r: > 0.1 small, > 0.3 medium, and > 0.5 large effect).

IQR=0.11 cm). The median distance per condition for the head-mounted group in ascending order are: *paper* instructions (Md=0.04 cm, IQR=0.09 cm) and *head-mounted* instructions (Md=0.04 cm, IQR=0.10 cm). The distances are compared in Figure 6.

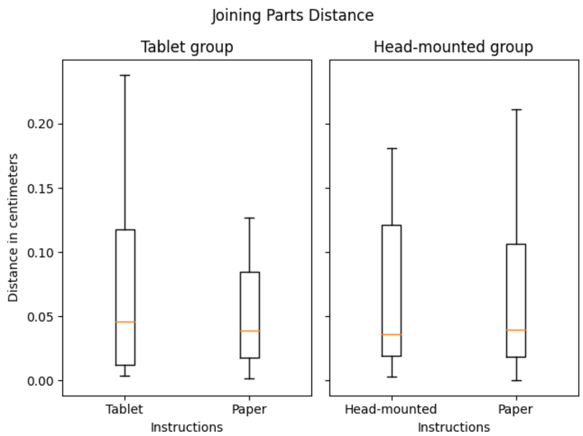


Figure 6: The median joining parts distance per group without outliers.

A Shapiro-Wilk-Test showed that our data are not normally distributed ($p < 0.001$). A Wilcoxon Signed-rank test showed no significant differences between the distances in the tablet group (W_{min}=380.0, z=-0.403, p=0.686, r=0.064) as in the head-mounted group (W_{min}=394.0, z=-0.215, p=0.831, r=0.034). A Mann-Whitney U-test also showed no significant difference between the distances of the *tablet* and *head-mounted* instructions (U_{min}=773.0, z=0.255, p=0.799, r=0.040). A Mann-Whitney U-test showed no significant difference between the distances of the *paper* instructions of both groups (U_{min}=753.0, z=0.447, p=0.655, r=0.071).

Total Duration. We consider the total duration of our four steps of the adhesive bonding process per condition. The median duration in seconds for the tablet group in ascending order are: *tablet* instructions (Md=179.79 s, IQR=76.84 s) and *paper* instructions (Md=292.33 s, IQR=111.79 s). The median duration in seconds for the head-mounted group in ascending order are: *head-mounted* instructions (Md=183.11 s, IQR=
93.13 s) and *paper* instructions (Md=229.62 s, IQR=174.94 s). The process times for each group are compared in Figure 7.

A Shapiro-Wilk-Test showed that our data of the tablet group is normally distributed ($p > 0.10$). Therefore, we used a two-tailed paired T-test for normally distributed data and parametric, related groups. This showed no significant difference between the duration in the tablet group for $p < 0.05$ (M=112.83 s, S_M=51.39, t(9)=1.833, t=-2.195, p=0.056, d_z=0.694). A Wilcoxon Signed-rank test showed no significant differences between the duration in the head-mounted group (W_{min}=13.0, z=-1.478, p=0.139, r=0.467). A Mann-Whitney U-test also showed no significant difference between the duration of the *tablet* and *head-mounted* instructions (U_{min}=46.0, z=0.265, p=0.791, r=0.084). A Mann-Whitney U-test showed no significant difference between the distances of the *paper* instructions of both groups (U_{min}=43.0, z=0.491, p=0.623, r=0.155).

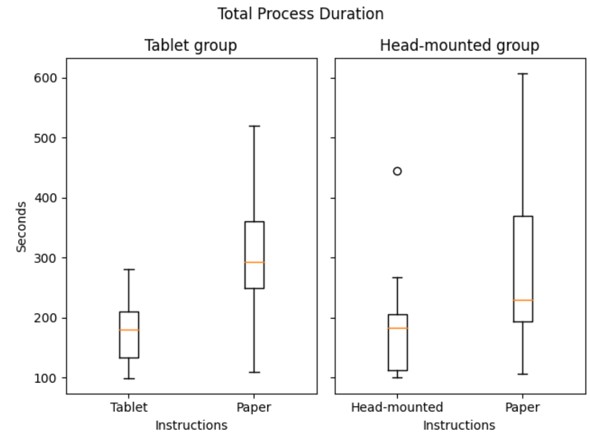


Figure 7: The median duration per group.

Timer. In order to check whether the participants understood the two different types of timers, we measured the number of correctly observed timers (see Figure 8).

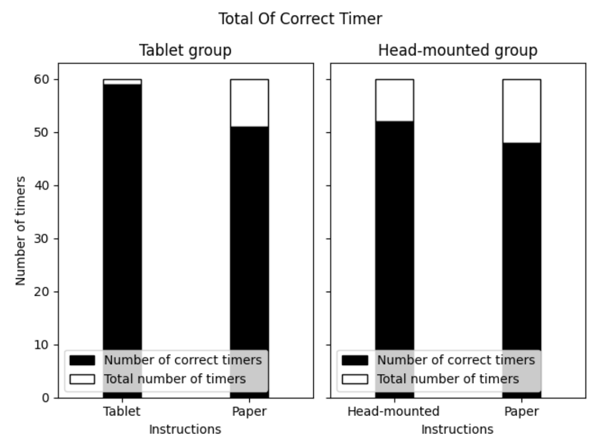


Figure 8: The number of correct timers in total.

The median number of correct timer per condition for the tablet group in ascending order are: *paper* instructions (Md=5.00, IQR=1.00) and *tablet* instructions (Md=6.00, IQR=0.00). The median number of correct timer per condition for the head-mounted group in ascending order are: *head-mounted* instructions (Md=6.00, IQR=0.75) and *paper* instructions (Md=6.00, IQR=1.75).

A Shapiro-Wilk-Test showed that our data are not normally distributed ($p < 0.001$). We used a Sign test for not normally distributed and non-parametric, related groups because the difference score of a subject was often zero under both conditions[10] for a Wilcoxon Signed-rank test. We found no significant difference between the number of correct timers in the tablet group (X=2.5, p=0.063) as in the head-mounted group (X=0.5, p=1.0) for $p<0.05$. A Mann-Whitney U-test showed no significant difference between

[10]If the difference score is often zero, the sample size in the Wilcoxon Signed Rank test is reduced and undermines the reliability of this test. To be able to make a statement nevertheless, the Sign test was applied to non-normally distributed, related groups.

the number of correct timers of the *tablet* and *head-mounted* instructions (U_{min}=39.0, z=1.136, p=0.256, r=0.359). A Mann-Whitney U-test showed no significant difference between the distances of the *paper* instructions of both groups (U_{min}=46.5, z=0.245, p=0.807, r=0.077).

System Usability Scale. With the help of the System Usability Scale (SUS) [5] we were able to evaluate the usability of the different instructions. The usability was best rated for the *tablet* instructions, followed by the *head-mounted* instructions and the *paper* instructions. In contrast, the *head-mounted* and *paper* instructions hardly differ. The median score for the tablet group in ascending order are: *paper* instructions (Md=58.75, IQR=23.13) and *tablet* instructions (Md=90.00, IQR=4.38). The median score for the head-mounted group in ascending order are: *paper* instructions (Md=66.25, IQR=17.50) and *head-mounted* instructions (Md=66.25, IQR=22.50). The score for each group are compared in Figure 9.

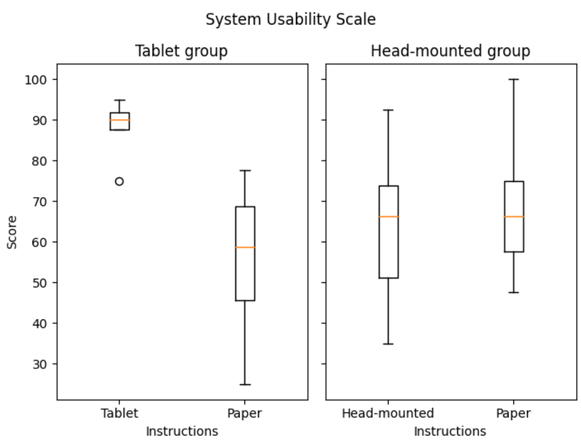


Figure 9: The median score per group.

Since these are ordinal data, we directly used the Wilcoxon Signed-rank test and Mann-Whitney U-test. A Wilcoxon Signed-rank test showed a significant difference between the score in the tablet group (W_{min}=0, z=-2.803, p=0.005, r=0.886). A Wilcoxon Signed-rank test showed no significant difference between the score in the head-mounted group (W_{min}=16.5, z=-0.715, p=0.475, r=0.226). A Mann-Whitney U-test showed a significant difference between the score of the *tablet* and *head-mounted* instructions (U_{min}=11.0, z=2.932, p=0.003, r=0.927). A Mann-Whitney U-test showed no significant difference between the score of the *paper* instructions of both groups (U_{min}=33.0, z=1.251, p=0.211, r=0.396).

Task Load. The results of the task load ratings as measured by the raw NASA Task Load Index (TLX) [10] are shown in Table 3. The *tablet* instructions induced the lowest mental, physical and temporal demand, as well as the best perceived performance and the lowest effort and frustration compared to all other instructions. The *head-mounted* instructions, in turn, induced the least mental and temporal stress, as well as the least effort compared to the *paper* instructions in the head-mounted group. In this group the *paper* instructions follow with the lowest physical demand, as well as the lowest frustration. Between the *tablet* and *paper* instructions (tablet group), we found a significant difference for the scales mental demand, performance, effort and frustration. Since these are ordinal data, we directly used the Wilcoxon Signed-rank test (see Table 2).

Table 2: Significant differences for the scales of the TLX in the tablet group (r: > 0.1 small, > 0.3 medium, and > 0.5 large effect).

Scale	W_{min}	z	p	r
Mental Demand	28.0	-2.414	0.016	0.763
Performance	1.5	-2.120	0.034	0.671
Effort	36.0	-2.529	0.011	0.799
Frustration	27.0	-2.209	0.027	0.698

In the head-mounted group, between the *tablet* and *head-mounted* instructions and between the *paper* instructions of both groups, we found no significant difference for each scale.

Subjective Measures. After each condition, we asked participants to answer one questions with a 5-point Likert-item (1=strongly disagree, 5=strongly agree). The results are shown in Figure 10. In the tablet group, participants were neutral for the *paper* instructions (Md=3.00, IQR=1.75), while almost all stated that the *tablet* instructions (Md=5.00, IQR=0.00) strongly supported them. In the head-mounted group, the participants stated that the *head-mounted* instructions (Md=4.00, IQR=1.00) and the *paper* instructions (Md=4.00, IQR=1.00) supported them equally. Because of the ordinal data no test for normal distribution is necessary. A Wilcoxon Signed-rank test showed a significant difference between the question in the tablet group (W_{min}=2.0, z=-2.257, p=0.024, r=0.714). A Wilcoxon Signed-rank test showed no significant difference between the question in the head-mounted group (W_{min}=7.0, z=-1.265, p=0.206, r=0.422). A Mann-Whitney U-test showed no significant difference between the *tablet* and *head-mounted* instructions (U_{min}=26.0, z=1.701, p=0.089, r=0.567). A Mann-Whitney U-test showed no significant difference between the *paper* instructions of both groups (U_{min}=32.0, z=1.073, p=0.283, r=0.358).

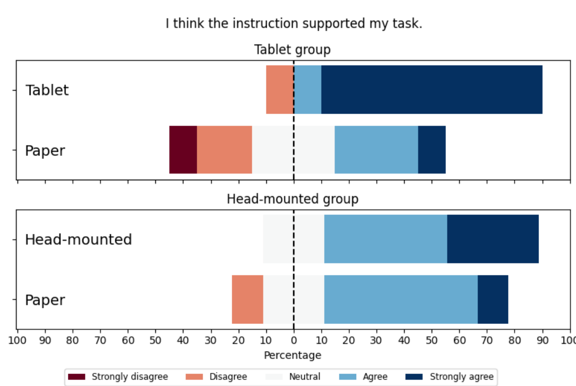


Figure 10: Results from Likert-item questionnaire.

Table 3: Task load ratings (Raw-TLX) for all instructions (values range from 1 (very low) to 10 (very high)).

Instructions	**Mental Demand**	**Physical Demand**	**Temporal Demand**	**Performance**	**Effort**	**Frustration**
	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)	Median (IQR)
Tablet	**2.00 (1.75)**	**1.00 (0.75)**	**2.00 (2.75)**	**9.00 (1.75)**	**2.00 (1.00)**	**1.00 (1.00)**
Paper	4.00 (4.25)	1.50 (1.00)	3.00 (1.00)	7.50 (2.50)	4.00 (3.00)	4.50 (5.00)
Head-mounted	**2.50 (1.00)**	2.00 (2.00)	**2.50 (1.75)**	7.50 (1.75)	**3.00 (1.00)**	3.50 (2.75)
Paper	4.50 (4.25)	**1.00 (1.00)**	4.00 (0.00)	**7.50 (2.50)**	3.50 (2.00)	**3.00 (2.75)**

6 DISCUSSION

Position Distance. We found that the *AR* instructions differed from the *paper* instructions in the correct positioning of the spacers and joining parts. In this context, we found a significant difference between the *tablet* and *head-mounted* instructions in the positioning of the spacers. We think that the participants first had to get used to the *head-mounted* instructions in the first step. In this study, we can partially confirm our hypothesis H_1, since there was a positive but no significant difference between the *AR* instructions and the *paper* instructions.

Total Duration. The process duration differed slightly but not significantly between the *tablet* and *paper* instructions. There was also no significant difference between the *head-mounted* and *paper* instructions, nor between the two *AR* instructions. In general, however, the process time was higher for the *paper* instructions. We suspect that reading and understanding the paper manual and starting the manual timers resulted in increased process times.

Timer Measurement. For the minimum timer (step 2), we measured whether the four joining parts had flashed off precisely 30 seconds. If the minimum time span of 30 seconds per part was observed, the respective timer was evaluated as correct. With the maximum timer (step 4) we checked whether the joining parts were bonded within a maximum time span of 20 seconds. Here it was measured how long each participant needed to join the parts. With a value 0<t<20 the timer was evaluated as correct. Otherwise it can be assumed that the timer was misunderstood and the bonding would fail in reality. Based on our measurement we could see that the timers were understood in all instructions. However, the *tablet* instructions caused the least amount of errors, which confirms our hypothesis H_1 at least for the tablet group.

System Usability Scale. The system usability scale is best rated for the *tablet* instructions, followed by the *head-mounted* and *paper* instructions. We found a significant difference in the tablet group, which confirms our hypothesis H_2. We assume that the participants already had experience in using mobile devices such as tablets and smartphones, so no learning effort was necessary. The *head-mounted* and *paper* instructions received the same median and hardly differ from each other.

Task Load. The results of the Raw TLX rating confirm the previous results (see Table 4). The *tablet* instructions are ahead in all ratings. We also found that the *tablet* instructions resulted in significantly lower mental demand, effort and frustration as in higher performance compared to the *paper* instructions. Therefore, the results confirm our hypothesis H_2. Here it is particularly interesting that the load is lower compared to both *head-mounted* and *paper* instructions, which is a further indication that the acceptance of mobile devices such as tablets is given and currently even preferred compared to traditional paper instructions and futuristic solutions such as head-mounted instructions.

Subjective Measures. The subjective measures showed that all of our instructions help the participants to understand their tasks. We found a significant better rating for the *tablet* instructions compared to the *paper* instructions, which confirms our hypothesis H_2. Overall, the *tablet* instructions received the best rating from the participants. In the head-mounted group, the *head-mounted* instructions were rated identical to the *paper* instructions. The participants stated that the head-mounted display felt uncomfortable. Glasses wearers in particular felt impaired. We think that these are possible reasons for the same rating. In addition, this information shows us that the paper manual was fairly constructed and is suitable as a baseline condition.

Distance measurement. To measure the correct positioning of spacer and joining parts within the (gluing) device, we have calculated the distance to the optimum position in each case (see section 5.2). This distance is a approximate value for us, which can vary by 0.2 cm after several previous measurements. This is due to the fact that the angle of view of the device through the camera of the tablet but especially of the Microsoft HoloLens causes a certain distortion. An exact measurement is therefore not possible due to the camera perspectives. But as we work with reference values that we determined before, we can get a valid estimation of the positioning by means of the measurements.

Marker detection. A challenge of our study was the marker detection, which is strongly dependent on environmental characteristics such as light conditions and camera focus. For this purpose, the markers have to meet certain requirements such as "rich in detail", "good contrast" and "no repetitive patterns" in order to be well detected. Another feature is the size of the markers. Our markers had a very small size due to the spacer and joining parts and we had to ensure recognition by placing boxes on the markers in the AR-View. So the participants could always be sure that the measurement was performed correctly.

Limitations. A technical limitation was the limited field of view of the Microsoft HoloLens. This means that the potential of human vision cannot be optimally used to naturally enrich visual perception through AR. However, by using the Microsoft HoloLens with their limited field of view, we were able to measure the performance of what is currently possible for head-mounted devices. For this purpose we classified different time variables such as time periods and points in time and implemented different scenarios. A limitation of the study design was the low participation and the selection of participants. The low participation could have caused the positive but not significant results between the comparison of the investigated conditions. In addition, the participants mainly represent students without any relation to adhesive bonding technology. A study with participants of a real training environment (like in section 3) for bonding technology training courses would perhaps provide more informative results. Although our participants were also beginners, they were not related to adhesive bonding technology. On the basis of the evaluation, it could also be determined that the *paper* instructions in the head-mounted group performed worse than in the tablet group. As the selection of the participants was random and without knowledge of persons, we cannot further explain this phenomenon.

7 CONCLUSION

When people are working in time-critical processes where the precise spatial positioning of parts also plays an important role, inexperienced workers, in particular, need understandable and straightforward instructions. In this way, process-related safety factors can be better observed, error rates reduced, and motivation factors increased. With our study, we show possibilities to instruct, especially beginners, safely in adhesive bonding processes. Thus, training can be implemented faster and safer, as, on the one hand, the hands are "kept free," and on the other hand, the time required for comprehension is reduced by the AR display. In this work, we compared two different AR assembly instructions with a paper-based variant while the worker could perform the various work steps without using his hands. We found that both the *tablet* and *head-mounted* instructions supported the participants in the process execution and resulted in reduced process times. The *paper* instructions required more training, which increased the overall execution time. Participants reported that they felt more comfortable with the *AR* instructions compared to the *paper* instructions. We explain this by the fact that we had inexperienced participants who were confronted with the adhesive bonding technique for the first time.

Our results also show that although a head-mounted AR device allows hands-free working, the technology here needs to progress even further in order to be accepted and used permanently. Especially glasses wearers have complained about comfort problems.

8 FUTURE WORK

For the future, we consider doing further experiments for time-dependent manual processes. In this context, (semi-)automatic process control using AR instructions is planned so that the system automatically adapts to the user's speed and learning curve. We assume that the user prefers the head-mounted solution due to the freedom of movement, stereoscopic vision, and advanced augmentation compared to the tablet solution for flexible workplaces. In addition, we want to analyze different time-related aspects with AR and examine how parallel times can best be presented to the user. For this purpose, we classify different time variables such as time periods and points in time and implement different scenarios.

REFERENCES

[1] J. Alves, B. Marques, M. Oliveira, T. Araújo, P. Dias, and B. S. Santos. 2019. Comparing Spatial and Mobile Augmented Reality for Guiding Assembling Procedures with Task Validation. In *2019 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. 1–6. https://doi.org/10.1109/ICARSC.2019.8733642

[2] Ronald T. Azuma. 1997. A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments* 6, 4 (1997), 355–385. https://doi.org/10.1162/pres.1997.6.4.355 arXiv:https://doi.org/10.1162/pres.1997.6.4.355

[3] Ó. Blanco-Novoa, T. M. FernáNdez-CaraméS, P. Fraga-Lamas, and M. A. Vilar-Montesinos. 2018. A Practical Evaluation of Commercial Industrial Augmented Reality Systems in an Industry 4.0 Shipyard. *IEEE Access* 6 (2018), 8201–8218. https://doi.org/10.1109/ACCESS.2018.2802699

[4] 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* (Island of Rhodes, Greece) *(PETRA '17)*. Association for Computing Machinery, New York, NY, USA, 75–82. https://doi.org/10.1145/3056540.3056547

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

[6] T. P. Caudell and D. W. Mizell. 1992. Augmented reality: an application of heads-up display technology to manual manufacturing processes. In *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*, Vol. ii. 659–669 vol.2. https://doi.org/10.1109/HICSS.1992.183317

[7] DIN 2304-1:2019-10 2019. Adhesive bonding technology - Quality requirements for adhesive bonding processes, Part 1: Adhesive bonding process chain.

[8] Markus Funk, Thomas Kosch, and Albrecht Schmidt. 2016. Interactive Worker Assistance: Comparing the Effects of in-Situ Projection, Head-Mounted Displays, Tablet, and Paper Instructions. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Heidelberg, Germany) *(UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 934–939. https://doi.org/10.1145/2971648.2971706

[9] G. Habenicht. 2009. *Kleben - erfolgreich und fehlerfrei: Handwerk, Praktiker, Ausbildung, Industrie*. Vieweg+Teubner Verlag. https://books.google.de/books?id=BskNkFjqDJsC

[10] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909 arXiv:https://doi.org/10.1177/154193120605000909

[11] Melynda Hoover. 2018. An evaluation of the Microsoft HoloLens for a manufacturing-guided assembly task. *Graduate Theses and Dissertations. 16378.* https://doi.org/10.31274/etd-180810-6008

[12] Lei Hou, Xiangyu Wang, and Martijn Truijens. 2015. Using Augmented Reality to Facilitate Piping Assembly: An Experiment-Based Evaluation. *Journal of Computing in Civil Engineering* 29, 1 (2015), 05014007. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000344 arXiv:https://ascelibrary.org/doi/pdf/10.1061/%28ASCE%29CP.1943-5487.0000344

[13] B. C. Kress and W. J. Cummings. 2017. 11-1: *Invited Paper* : Towards the Ultimate Mixed Reality Experience: HoloLens Display Architecture Choices. *SID Symposium Digest of Technical Papers* 48, 1 (May 2017), 127–131. https://doi.org/10.1002/sdtp.11586

[14] F. Loch, F. Quint, and I. Brishtel. 2016. Comparing Video and Augmented Reality Assistance in Manual Assembly. In *2016 12th International Conference on Intelligent Environments (IE)*. 147–150. https://doi.org/10.1109/IE.2016.31

[15] N. Pathomaree and S. Charoenseang. 2005. Augmented reality for skill transfer in assembly task. In *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005*. 500–504. https://doi.org/10.1109/ROMAN.2005.1513829

[16] Carole Plasson, Dominique Cunin, Yann Laurillau, and Laurence Nigay. 2019. Tabletop AR with HMD and Tablet: A Comparative Study for 3D Selection. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces* (Daejeon, Republic of Korea) *(ISS '19)*. Association for Computing Machinery, New York, NY, USA, 409–414. https://doi.org/10.1145/3343055.3360760

[17] Loukas Rentzos, Stergios Papanastasiou, Nikolaos Papakostas, and George Chryssolouris. 2013. Augmented Reality for Human-based Assembly: Using Product and Process Semantics. *IFAC Proceedings Volumes* 46, 15 (2013), 98 – 101. https://doi.org/10.3182/20130811-5-US-2037.00053 12th IFAC Symposium on Analysis, Design, and Evaluation of Human-Machine Systems.

[18] Eric Rose, David Breen, Klaus H. Ahlers, Chris Crampton, Mihran Tuceryan, Ross Whitaker, and Douglas Greer. 1995. 25 - Annotating Real-World Objects Using Augmented Reality. In *Computer Graphics*, Rae Earnshaw and John Vince (Eds.). Academic Press, Boston, 357 – 370. https://doi.org/10.1016/B978-0-12-227741-2.50029-3

[19] A. Sanna, F. Manuri, F. Lamberti, G. Paravati, and P. Pezzolla. 2015. Using handheld devices to support augmented reality-based maintenance and assembly tasks. In *2015 IEEE International Conference on Consumer Electronics (ICCE)*. 178–179. https://doi.org/10.1109/ICCE.2015.7066370

[20] B. Schwerdtfeger and G. Klinker. 2008. Supporting order picking with Augmented Reality. In *2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*. 91–94.

[21] Sule Serubugo, Denisa Skantarova, Lasse Kjærsgård Nielsen, and Martin Kraus. 2017. Comparison of Wearable Optical See-through and Handheld Devices as Platform for an Augmented Reality Museum Guide. *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications* (2017), 179–186. https://doi.org/10.5220/0006093901790186

[22] Anna Syberfeldt, Magnus Holm, Oscar Danielsson, Lihui Wang, and Rodney Lindgren Brewster. 2016. Support Systems on the Industrial Shop-floors of the Future – Operators' Perspective on Augmented Reality. *Procedia CIRP* 44 (2016), 108 – 113. https://doi.org/10.1016/j.procir.2016.02.017 6th CIRP Conference on Assembly Technologies and Systems (CATS).

[23] Arthur Tang, Charles Owen, Frank Biocca, and Weimin Mou. [n.d.]. Comparative effectiveness of augmented reality in object assembly. In *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2003), Ft.* 73–80.

[24] Stefan Wiedenmaier, Olaf Oehme, Ludger Schmidt, and Holger Luczak. 2003. Augmented Reality (AR) for Assembly Processes Design and Experimental Evaluation. *International Journal of Human–Computer Interaction* 16, 3 (2003), 497–514. https://doi.org/10.1207/S15327590IJHC1603_7 arXiv:https://doi.org/10.1207/S15327590IJHC1603_7

[25] Xianjun Sam Zheng, Cedric Foucault, Patrik Matos da Silva, Siddharth Dasari, Tao Yang, and Stuart Goose. 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* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 2125–2134. https://doi.org/10.1145/2702123.2702305

[26] Feng Zhou, Henry Been-Lirn Duh, and Mark Billinghurst. 2008. Trends in Augmented Reality Tracking, Interaction and Display: A Review of Ten Years of ISMAR. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '08)*. IEEE Computer Society, Washington, DC, USA, 193–202. https://doi.org/10.1109/ISMAR.2008.4637362