In-Situ Instructions Exceed Side-by-Side Instructions in Augmented Reality Assisted Assembly

Jonas Blattgerste, Patrick Renner, Benjamin Strenge and Thies Pfeiffer

CITEC - Cluster of Excellence Cognitive Interaction Technology

Bielefeld University

{jblattgerste,prenner,bstrenge,tpfeiffer}@techfak.uni-bielefeld.de

ABSTRACT

Driven by endeavors towards Industry 4.0, there is increasing interest in augmented reality (AR) as an approach for assistance in areas like picking, assembly and maintenance. In this work our focus is on AR-based assistance in manual assembly. The design space for AR instructions in this context includes, e.g., side-by-side, 3D or projected 2D presentations. In previous research, the low quality of the AR devices available at the respective time had a significant impact on performance evaluations. Today, a proper and up-to-date comparison of different presentation approaches is missing.

This paper presents an improved 3D in-situ instruction and compares it to previously presented techniques. All instructions are implemented on up-to-date AR hardware, namely the Microsoft HoloLens. To support reproducible research, the comparison is made using a standardized benchmark scenario. The results show, contrary to previous research, that in-situ instructions on state-of-the-art AR glasses outperform side-by-side instructions in terms of errors made, task completion time, and perceived task load.

KEYWORDS

Augmented Reality; Assistance Systems; Head-Mounted Displays; Smart Glasses; Benchmarking.

1 INTRODUCTION

While the digital office can be considered the norm today in many businesses, digitalization has not yet reached classic industrial workspaces. Manual assembly tasks often still rely on paper instructions. However, within the context of Industry 4.0 [8] and in particular the goal of a highly variable manufacturing with a batch size of one, a change towards using modern technologies is at stake. Where automation is not possible, assistive systems can help workers to maintain or even increase their productivity under this increased task load. They can moreover help integrating and training people with cognitive disabilities [16] by giving instructions and monitoring, e.g., the state of assembly.

Augmented reality is capable of projecting assistive instructions in-view of the user or directly in-situ at the object-of-interest. Projections can be in 2D or in 3D, e.g., when using binocular AR glasses. A variety of different AR devices (projectors, tablets, monocular and binocular glasses) and approaches for giving instructions have been presented. In the past, many evaluations compared AR instructions separately to classic - e.g. paper-based - instructions. Past mobile AR devices thereby often had performance issues leading to slow and instable tracking [14, 26] or a low visual fidelity due to low-resolution displays [13]. Only recently mobile AR devices with reasonable processing power, tracking accuracy and latency, as well as display quality were released to the market, e.g. Microsoft HoloLens.

In this work, our aim is to evaluate assembly instructions on AR glasses. In past experiments these performed worse than projection-based approaches [5], but there was already some evidence that approaches on modern AR glasses may perform similarly well with the advantage of not being bound to a specific workspace [4]. Previous comparisons also indicated that side-by-side instructions outperform in-situ instructions using AR glasses [14]. However, we anticipate that these results do no longer hold in light of modern hardware. Therefore, in this paper we focus on the comparison of different in-situ and side-by-side instructions for an assembly task which were implemented on a Microsoft HoloLens. In order to generate reproducible results, we chose a standardized study scenario, originally proposed by Funk et al. [6], where a pick-and-assemble task using LEGO DUPLO bricks has to be solved by study participants.

We evaluated four different AR instruction techniques for placing a brick in this scenario: An improved 3D in-situ visualization of the brick matching the exact target position, orientation, and color; a 2D in-situ visualization simulating a projected instruction; an animated wireframe representation of the brick, which should avoid occlusions of the target assembly position; and a side-by-side instruction based on the approach proposed by Khuong et al. [14], which contains all previously assembled bricks.

2 RELATED WORK

Assistive systems have recently reached various areas of Industry 4.0. The continuing development of AR hardware bears a huge potential for increasing productivity.

2.1 Picking Tasks

Picking tasks are a classic application area for AR assistance. Using AR glasses, the picker can be guided towards the target shelf in

a hands-free manner. Schwerdtfeger et al. [24] visualize either a compass combined with a rubber line leading to the target drawer or a tunnel showing the way towards it. Reif et al. [20] compare a similar approach using a tunnel of circles as well as a contour highlight at the drawer to classic paper-based guidance. Another evaluation was conducted by Guo et al. [9]: They compare HUD information on AR glasses and a display mounted on a cart to classic picking lights (small lamps at each drawer indicating the relevant drawers to pick from). The results of these evaluations suggest that AR-based picking guidance outperforms classic techniques.

For more complex environments, such as industry halls with a number of shelves to pick from, a combination of techniques can be useful. Henderson et al. [12] propose a combination of 2D and 3D arrows for huge orientation changes, highlights when targets come into the AR FOV and finally fading out all visualizations in order to prevent them from occluding targets. In the work presented in this paper, the picking subtask comprises a small set of options, all within reaching distance and in the field-of-view of the participant. We thus followed the advice and only used unobtrusive highlights to guide the user towards the correct picking target.

2.2 Assembly Tasks

In contrast to guiding techniques for picking tasks, in the area of assembly the focus lies on how to conduct actions in a correct way. This begins with choosing the correct objects (materials or tools) in the workspace, which can be supported by highlighting them [3]. Instructions how to conduct an action can, e.g., be provided using video recordings played in the user's FOV [18]. Petersen et al. [19] extend this approach by in-situ projections of previously recorded videos directly as an overlay to the current situation. They moreover analyze the current status and provide feedback by colorizing the user's hands.

Pure virtual instructions can be provided in various ways. A straight-forward approach is to project them directly into the workspace at the task relevant location [17]. This approach has been shown to support product assembly without prior knowledge [22]. A comparison of in-view instructions on AR glasses with tablet or paper instructions revealed that only visualizing the paper instructions on mobile devices does not improve performance of workers [7].

To improve over static instructions, Gupta et al. [10] track the orientation of the current part and show instructions on a monitor matching the part's orientation. Their system outperformed traditional paper instructions in terms of speed and accuracy. Using AR glasses, in-situ instructions can be given in 3D: Alvarez et al. [1] used augmented in-situ instructions for speeding up disassembly tasks. Objects to disassemble can be highlighted and an arrow shows the direction of the action to conduct. Stanimirovic et al. [25] use a tablet computer for showing in-situ visualizations in form of superimposed 3D animations and additional billboarded textual descriptions. Also in a LEGO DUPLO assembly scenario, Khuong et al. [14] proposed a side-by-side visualization for instructing workers. Compared to a wireframe-based in-situ instruction, their results show that spatial separation of the real and the virtual model performed better than the target-aligned in-situ variant. They argue that in-situ projections are sensitive to misalignment, latency and conflicting depth cues.

However, they state that overlaid in-situ instructions could reduce spatial ambiguity when those issues were solved.

This is the starting point of the work presented here: Our claim is that current AR glasses have overcome latency and accuracy issues in an extent that in-situ instructions should actually perform best in terms of speed and accuracy compared to both other instruction techniques on AR glasses and instructions on other devices. First evidence supporting this claim has been found in our prior work [4]. The 3D in-situ visualization used there, however, was not optimal, as AR objects sometimes occluded already placed real objects, which made the disambiguation of the correct placement difficult. In the improved version presented in this paper, occlusions of AR objects by real objects are taken into account. To substantiate our claim, we implemented AR instructions similar to the ones proposed by Khuong et al. [14] and the improved 3D in-situ instruction in our evaluation.

2.3 Projected and Mobile Augmented Reality

Assistive systems based on AR not only differ in their user interface design, but also in the device used for presentation. Projected AR led to promising results concerning task performance [22, 23] in the past. Korn et al. [15–17] used projected AR to help people with cognitive disabilities in the workplace by displaying in-situ instructions and monitoring the current assembly state. Rodriguez et al. [21] even used projection mapping for visualizing information correctly overlaid to objects in the workspace. An evaluation conducted by Funk et al. [7] suggests that projected in-situ AR performs better than in-view instructions using either an HMD, a tablet or paper-based instructions. However, projected AR has the disadvantage of being bound to a specific workspace. It cannot move with the worker like AR glasses do. Moreover, projections are limited to 2D. In our evaluation, we focus on AR glasses especially because of the advantage of 3D visualizations. To learn about the impact of 2D vs. 3D visualizations, we also evaluate a 2D in-situ instruction.

2.4 Evaluation Scenario

In the literature, mostly two different scenarios for evaluating assistance in assembly tasks were chosen: For maximizing realism, workbench scenarios are used (e.g. in [1, 2]) or specific areas of application like furniture assembly [11] or notebook maintenance [19]. The other range of scenarios abstracts from real tasks, e.g., by using LEGO bricks [10, 14, 17]. Funk et al. [6] proposed a standardized scenario for comparing different instruction techniques. It encompasses a picking part as well as a placing part using LEGO DUPLO bricks. In that scenario, already a number of studies was conducted [4, 7]. We adopted this scenario to support the work on standardized procedures for benchmarking different AR instruction designs in a reproducible way.

3 IMPLEMENTATION OF THE INSTRUCTION TECHNIQUES

The assembly environment introduced by Funk et al. [6] can be split into two areas: The spare part area containing eight blue container bins and the assembly area with a green 24x24 Lego Duplo plate. Therefore the needed AR instructions can also be split into two
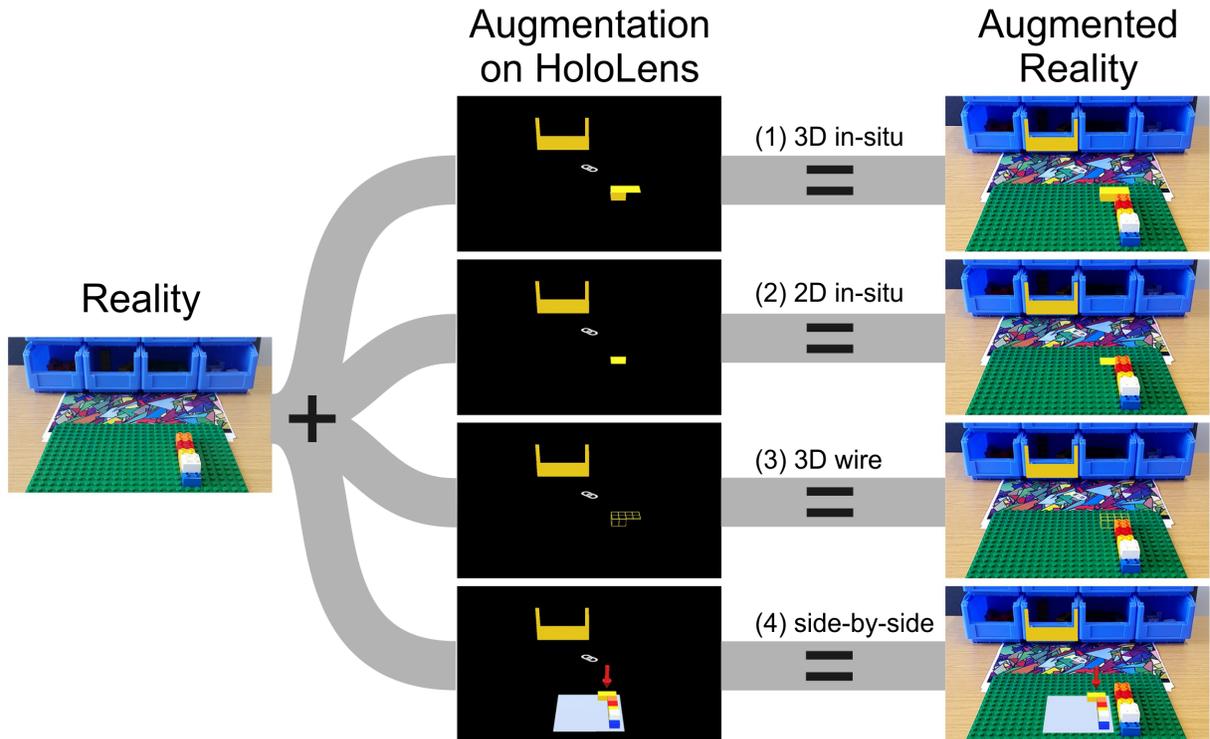
Figure 1: The standardized assembly scenario (Reality), the computerized augmentations on the Microsoft HoloLens (Augmentation) and the result (Augmented Reality). The Augmentations (from top to bottom): 3D in-situ instructions (1), 2D in-situ instructions (2), 3D animated wireframe instructions (3) and side-by-side instructions (4)

categories: Picking instructions (for the container bins) and assembly instructions (on the Lego Duplo plate).

For the picking instructions Funk et al. [7] and Blattgerste et al. [4] both highlighted the corresponding container of the current brick of the assembly task. While Funk et al. [7] highlighted the top of the entrance of the container bin, in [4] we used a cross-hair to point to the correct container bin. We were able to show that this significantly reduced errors while picking a brick. Therefore we use a similar approach in this experiment, but instead of a cross-hair, which could possibly mask the entrance and make it harder to see the bricks inside the container bins, we decided to highlight the whole front side of the container in the corresponding color. This scheme was used for picking support in each evaluated instruction approach since we are mainly interested in the instructions for the assembly phase.

Similar to our previous experiment [4], we place an AR marker between both areas for the spatial registration between the physical space and the AR models. A visual anchor (chain-symbol) is visualized in AR directly overlapping to the physically counterpart to allow the user a visual estimation of registration precision.

For the AR assembly instructions we compare the following four approaches:

3.1 3D in-situ Instructions

The static 3D in-situ instructions (Figure 1.1) mainly follow the approach we introduced in [4]. A basic cuboid with the size and color of the corresponding brick is displayed at the exact target location of the current step of the task. We continue to argue that displaying an actual model of a LEGO DUPLO brick would not give any advantages over a plain cuboid in the scenario, it would only increase visual clutter. The results of the previous experiment showed that participants had problems with the interpretation of the visualization, as it did not consider the occlusion of the new block by previously placed (real) bricks. In the discussion, we argued that this problem, at least partially, could have been the reason for the high amount of errors we observed during the assembly phase [4] (0.92 errors per participant and task, resulting in a 2.875% chance for an error per placement). In the study at hand, we thus implemented a correct handling of occlusions to create a proper visualization of the expected target state.

3.2 2D in-situ Instructions

The static 2D in-situ instructions (Figure 1.2) simulate the projected AR introduced by Funk et al. [7], where the position of the current step of the assembly task is displayed as a highlight in a color matching that of the target block and not as an actual 3D object. Similar to

the 3D in-situ instructions, occlusion by previously placed bricks are taken into account, as this would have been the case in real projected AR, too. However, we did not explicitly implement false positives, i.e. the overlapping of projected target positions with already placed bricks at higher positions in off-center positions. These effects might happen when projections do not correctly compensate for the 3D structures currently on the workplace.

3.3 Animated 3D in-situ Wireframe

The 3D in-situ instructions with the animated wireframe (Figure 1.3) are based on the in-situ wireframe approach that was introduced by Khuong et al. [14]. Similar to their implementation, we use a wireframe with the corresponding size and color of the brick. We furthermore also attach an animation to it to suggest the snapping-in of that brick during assembly. Beside the animated wireframe of the current step, Khuong et al. [14] also displayed the previous layer of bricks as a static wireframe because they believed that this has the potential to reduce spatial ambiguity. In our implementation, we only display the animated wireframe of the current step as we believe AR tracking is already stable enough, therefore those problems should no longer occur and this would only add unnecessary clutter. As the assembly task that was used with their implementation was built upwards, they did not have potential occlusion of previously placed bricks. We therefore advance their implementation by adding the occlusion of previously placed bricks similar to the 3D and 2D in-situ instructions.

3.4 Side-by-side Instructions

The side-by-side in-situ instructions (Figure 1.4) also follow an implementation approach of Khuong et al. [14] where the instructions are not displayed directly aligned with but separately beside the assembly task. Furthermore not only the current brick of the task is displayed but everything that was build up to this point. While they displayed the instructions diagonally to the assembled figure with the same size, our implementation slightly differs. As the assembly task introduced by Funk et al. [6] is larger and shaped like an L, we had to display the instructions smaller (factor: 0.6) to fit them inside the L-shape for a fair comparison. If, e.g., the side-by-side instructions would have been placed with their original size at the same position, they would have occluded the actual assembled figure; if they would have been placed strictly beside the whole figure, an unfair disadvantage for this approach would have been introduced as additional head movement would have been necessary to get the visualization covered by the AR display. As every approach beside this one was an in-situ approach, we also displayed the side-by-side instructions in-situ but mounted them on a differently colored plate so participants would not get confused and try to place the bricks directly onto the displayed absolute position.

Feedback from participants in a non-representative preliminary study strongly suggested that an animation on the side-by-side instructions had "no benefit and was unnecessary" as it did not help to make the instructions more clear but introduced additional clutter. Therefore we furthermore altered the side-by-side instructions by replacing the animation of the current brick with an arrow. Using arrows for showing AR instructions already proved to be an intuitive approach in previous research [1, 12].

4 EXPECTATIONS

In contrast to previous work we believe that with the improved tracking of modern AR glasses like the Microsoft HoloLens, in-situ AR instructions should outperform side-by-side AR instructions in assembly tasks regarding task completion time, errors made and task load. Furthermore, when asked to rank the different approaches, we believe that participants correctly perceive in-situ instructions as being faster simply because the instructions are directly projected onto the actual assembly position. Moreover, we expect participants to falsely perceive side-by-side instructions as being more accurate, as the full assembly status up to the current brick is visualized and is always consistent (e.g. no tracking latencies/swimming). We also believe that, while all in-situ instructions should perform equally in terms of task completion time, 3D in-situ instructions should outperform 2D in-situ instructions with regard to errors made and task load.

5 METHODOLOGY

The experiment was conducted as a repeated measures experiment with the independent variable being the four approaches of AR instructions for placing bricks in assembly tasks. The dependent variables were task completion time, NASA (Raw) TLX score, errors made, and subjective rating of the instructions by the participants.

To make the results comparable we used the exact same 32-brick assembly task of [6] for every condition. To counterbalance possible bias due to order effects and especially learning effects, we used each permutation of the 4 AR instructions in one trial, thus each of the N=24 participants tested the instructions in different order.

To collect qualitative results, we asked the participants to fill out the NASA TLX questionnaire after each condition. After completion of all conditions, we asked the participants to rank the AR instructions with regard to speed, accuracy, learnability, usefulness, joy of use and preference.

For the quantitative results we filmed the participants while completing each assembly task, which allowed us to annotate the errors participants made. The system also recorded the time participants needed to place one brick by logging the time difference between two clicks of the HoloLens Clicker, as explained later.

5.1 Apparatus

The apparatus (Figure 3) of the experiment consisted of a table containing the assembly environment, a chair, a 360 degrees camera, the Microsoft HoloLens and an explanation sheet.

The assembly environment closely followed the standardized LEGO DUPLO assembly task introduced by Funk et al. [6] and therefore consisted of two areas, the spare part area containing 8 blue container bins where bricks in different colors were stored, and the assembly area with a green 24x24 LEGO DUPLO plate. Similar to our previous work [4], an AR marker was added between those areas for the initial registration of the tracking. The container bins, the green plate and the AR marker were all fixed to the table. The Microsoft HoloLens was the AR glasses of choice for this experiment, as it provides stable tracking at low latencies. The AR marker was only used to initially register the physical assembly environment with the virtual model. Then this registration was passed to the "spatial tracking" of the HoloLens, which was exclusively used
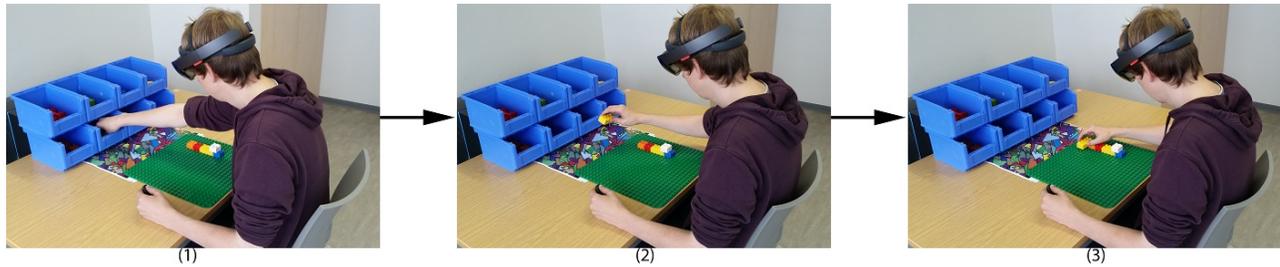
Figure 2: The procedure of a participant searching for the correct container and grabbing a brick (1), searching for correct assemble position (2) and assembling the brick (3)

Figure 3: The apparatus consisting of the LEGO DUPLO assembly task that was used for the experiment and the 360 degrees camera.

for tracking from then on. While the HoloLens still has technical limitations, such as the small FOV of 30 x 17 degrees, it allows for AR tracking without any accuracy issues in a sitting experience like in this experiment.

Furthermore, the 360 degree camera was placed to the left of the assembly environment to catch every hand movement of the participants and especially also the inside of the container bins. The explanation sheet for the AR instructions was placed to the right of the container bins.

5.2 Procedure

First, participants were asked to sign a declaration of consent and fill out a demographic questionnaire. After completing those, they were asked to sit down at the assembly environment. They were given a general explanation of the assembly environment and their impending task. Then they were introduced to the HoloLens and the 4 different AR instruction approaches were explained with the help of an explanation sheet that also contained a graphic for each AR instruction approach.

After communicating that they understood the experiment and the instructions, the HoloLens calibration tool was started and the participants were asked to put on the Microsoft HoloLens.

After calibrating the device, the application for the experiment was started. Participants were asked to look at the green LEGO DU-PLO plate and assemble the two red bricks that were now displayed as a part of the calibration-process. After placing those two bricks they were asked if the AR instructions and the real brick perfectly align. If they didn't, the experimenter was able to remotely correct the offset using a bluetooth keyboard.

The participants where then introduced to the HoloLens Clicker, a small handheld device with one button, that was used to forward to the next step of the task. They were explained that they can solve the task with either the right or the left hand and are supposed to hold the Clicker in the opposite one. Furthermore, they were instructed that they were supposed to press the Clicker once initially for the first brick and then always at the exact moment they released a brick from their hands after placing it. They were told that the first priority of the experiment was to finish the task without errors and the second priority to do this as fast as possible.

Participants then started with one of the four AR instruction approaches. First they were asked to solve a test task containing 8 steps to get used to the currently displayed AR instruction approach. Afterwards the 360 degrees camera was started and participants solved the 32 step assembly task (Figure 2) with the same AR instruction approach. After completing the task, the camera was stopped and they were handed the associated NASA TLX questionnaire. This was repeated for the three remaining AR instruction approaches.

After completing all conditions, they were finally handed the questionnaire that asked them to rank the approaches and then gave them the opportunity to write down comments, observations or suggestions regarding the experiment.

5.3 Participants

We conducted the experiment with 24 participants. The participants were aged between 18 and 31 (average = 23.72, SD = 3.17) and 13 of them were female. Most participants were students of our university.

6 RESULTS

As objective measures we recorded the task completion times and the errors participants made. Moreover, we let the participants report about their task load using the NASA TLX questionnaire after every condition. Finally, a ranking questionnaire had to be filled out. In the following, we report our measurements.
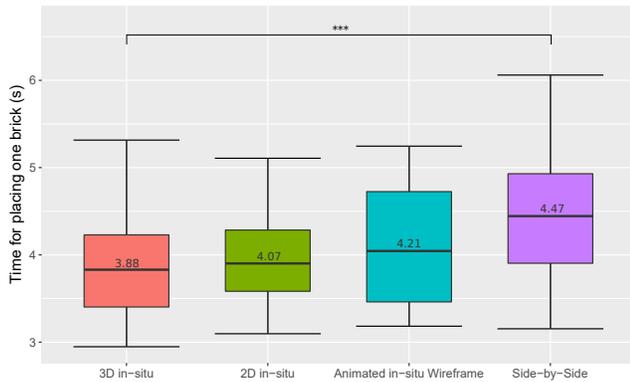
Figure 4: The time participants needed from picking a brick until placing it, using the different instruction techniques.

Figure 5: The percentage of erroneous attempts participants made in placing bricks. Error bars depict the standard error of the mean.

The completion times, the number of errors and the TLX scores have been statistically compared between the four instruction techniques using a one-way repeated measures ANOVA. When the ANOVA showed a significant difference between systems, pairwise comparisons have been conducted using paired t-tests with Bonferroni-Holm correction.

6.1 Task Completion Times

As we are interested only in the time needed to place a brick and thus kept picking assistance the same in all conditions, it is not necessary to divide task completion time in four different phases (t_{locate}, t_{pick}, t_{locate_pos}, $t_{assemble}$) as done in Funk et al. [6]. The first two phases measure the time until a brick is picked. In our case, we do not consider these. Therefore, we focus on the time spent on locating the correct assembly position. Participants clicked whenever having placed a brick. The time between two clicks thus encompasses all four phases, but as the picking guidance is constant over all conditions, only t_{locate_pos} and $t_{assemble}$ are supposed to differ due to the instruction techniques. We suppose that placing the brick ($t_{assemble}$) thereby is not strongly influenced by the instruction technique.

Figure 4 shows that the 3D in-situ instructions led to the fastest time for placing a brick. Participants needed on average 3.88 s (SD=1.04 s) using this technique. The side-by-side technique required most time to place a brick with a mean of 4.47 s (SD=2.27 s). In between, the 2D in-situ instructions lead to a task-completion-time of 4.07 s (SD=1.40 s) and using the animated in-situ wireframe, participants needed 4.21 s (SD=3.42 s). An ANOVA showed a significant difference between the systems, $p < .01$. The post-hoc tests revealed that the 3D in-situ required significantly less time for placing a brick than the side-by-side instructions ($p < .005$).

6.2 Average Number of Errors

The errors while assembling the bricks were counted. Overall, errors occurred seldom (see Figure 5): On average, participants made errors in 1.17 % (SD=2.37 %) of all brick placements. Being instructed by the 3D in-situ technique, only one participant made a single
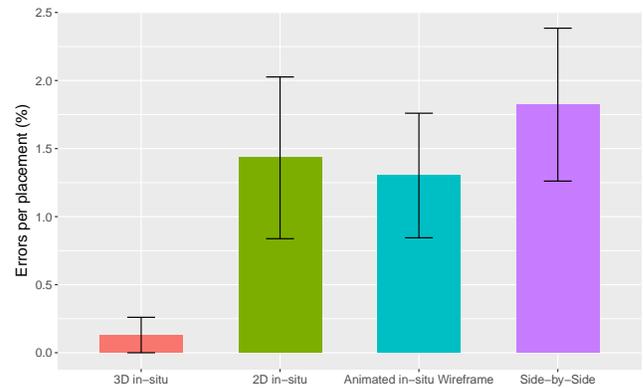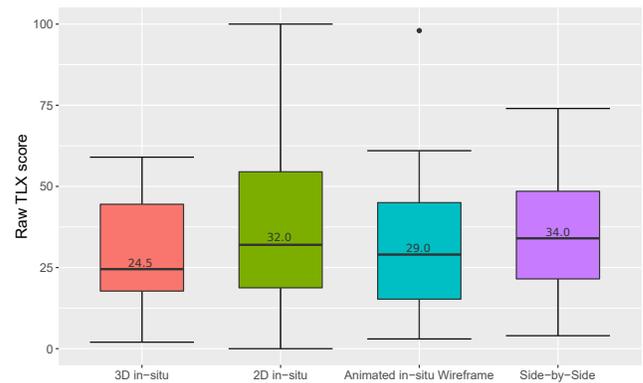

Figure 6: The task load participants reported when being instructed by the different techniques (NASA-TLX).

error at all, which equals 0.13 % overall. Using the other techniques, more errors were made. The animated in-situ wireframe instruction lead to 1.3 % errors (SD=2.24 %), the 2D in-situ instruction lead to 1.43 % errors (SD=2.91 %). Most errors were made using the side-by-side instructions with an average of 1.82 % errors (SD=2.75 %). An ANOVA showed a significant difference between the systems, $p < .05$. However, the post-hoc tests only revealed trends that the 3D in-situ instructions could reduce errors being made, especially compared to the side-by-side technique ($p = .055$).

6.3 Task Load

As illustrated by Figure 6, participants reported the least task load when using the 3D in-situ instruction technique with a mean TLX score of 29.6 (SD=17.4). The animated wireframe lead to a slightly higher TLX score of 32.8 (SD=21.9), followed by the side-by-side instructions which had a mean TLX score of 36.7 (SD=20.0). Participants reported the highest task load when being instructed using the 2D in-situ instructions with a score of 37.6 (SD=24.4). However,

an ANOVA did not show significant differences between these task load values.

6.4 Rankings

After finishing the experiment, participants were asked to report their subjective preferences regarding the four instruction techniques by rank-ordering them according to different categories: Speed, accuracy, learnability, usefulness, joy of use, and overall preference.

Regarding subjective speed, 50% of the participants thought that the 3D in-situ instructions helped them most to solve the task quickly. The other half shared their votes equally among the other three instruction techniques. In terms of subjective accuracy, in contrast to the objective results 45.8% of the participants considered the side-by-side instructions as most accurate, followed by the 3D in-situ instructions with 29.2%. The in-situ wireframe instructions were rated as most accurate by only 16.7%, the 2D in-situ instructions by 8.3%. At the other end, most often rated as least accurate were the side-by-side instructions (37.5% of the participants) and the 2D in-situ instructions (29.2%). Most participants regarded the 3D in-situ instructions (41.7%) and the side-by-side instructions (33.3%) as best learnable, followed by the wireframe in-situ instructions (20.8%), while only 4.2% of the participants considered the 2D in-situ instructions as easiest to learn.

Since the ordering can be considered as a forced-choice situation, we applied Fisher's exact test to analyze the respective contingency tables. Significant associations between instruction techniques and rankings were found regarding speed ($p < .01$), accuracy ($p < .01$), learnability ($p < .05$) and joy of use ($p < .05$), but not regarding usefulness or overall preference. Pairwise post-hoc tests using Wilcoxon signed-rank tests with Bonferroni-Holm correction showed that our 3D in-situ instructions were rated significantly better than the 2D in-situ approach in the categories speed ($p < .05$), learnability ($p < .01$) and joy of use ($p < .05$). The animated in-situ wireframe also ranked significantly higher than the 2D approach regarding joy of use ($p < .05$). No other significant pairwise differences were found.

7 DISCUSSION

Both the objective and the subjective results strongly suggest that our main expectation was met: The 3D in-situ instruction technique required least time for participants to fulfill their task and they made least errors using it. Especially in comparison to the side-by-side instructions, this technique was significantly faster. Moreover, participants reported the lowest task load with the 3D in-situ technique. In contrast to the findings in Khuong et al. [14], in this evaluation, the side-by-side technique led to the slowest task completion time, most errors and high task load. Thus, we believe that we could meet the proposed criteria of Khuong et al. [14] for successfully using in-situ AR visualizations: We established a solid tracking using the Microsoft HoloLens and ensured a stable high frame rate. The correct handling of occlusions by real objects seems to improve the error rate, which went down from 2.875% per placement in our previous work [4] to only 0.13% in this study. This way, it was possible to benefit from the theoretical advantages of in-situ visualizations.

Regarding task completion time, the 2D in-situ instructions and the animated 3D wireframe instructions required slightly more time than the 3D in-situ technique. We believe that the 2D instructions make it more difficult to disambiguate a 3D position, which could also be the reason for higher task load. In case of the 3D wireframe, we could observe participants waiting for the animation to finish, leading to longer time for placing a brick.

In comparison to our previous experiment [4], the number of errors made by the participants using all the in-situ instructions is considerably lower. In [4] we stated that the high number of average errors results from not correctly handling occlusions by real objects in the visualizations. Thus we believe that handling occlusions is indeed a crucial factor for making in-situ visualizations unambiguous. Additionally, 2D in-situ instructions seem not to be sufficient to clearly reference a 3D position, leading to more errors. The wireframe technique might suffer from its animation: Participants could misinterpret the moving location of a visualized brick and thus tend to make more errors. The 3D in-situ instruction technique did not suffer from these issues, thus participants made nearly no errors using it.

Observing the subjective ratings participants made regarding the speed of the techniques, these are in line with the objective results and our expectations. However, the ratings concerning accuracy interestingly are basically divided into two groups: One that considered the 3D in-situ instructions as most accurate and the side-by-side instructions relatively inaccurate and nearly one half that regarded the side-by-side instructions as most accurate. Thus, our expectation that side-by-side instructions are falsely rated as very accurate can only partly be confirmed, nearly one third of the participants rated the 3D in-situ technique in line with the objective results.

8 CONCLUSION

In our evaluation, we compared four approaches for assembly assistance: Three different 3D in-situ instruction approaches, and a side-by-side instruction approach in a LEGO DUPLO assembly scenario. Our contribution is two-fold: Firstly, we have shown that – in contrast to findings in the past – in-situ instruction techniques using AR glasses outperform the side-by-side variant when ensuring a stable and accurate tracking and a high frame rate. In addition to these technical requirements, handling occlusions of the visualizations by real objects was identified as a crucial factor. This is in line with the assumptions of Khuong et al. [14], whose evaluation was in favor of side-by-side instructions but suffered from problems concerning tracking accuracy. Secondly, our comparison of different in-situ instruction techniques revealed that in terms of task completion time, errors and task load, a static 3D visualization performs better than both a 2D projection and animated 3D instructions using a wireframe instead of a solid geometry.

Building on these findings, in future experiments we plan to also evaluate the influence of FOV size of AR devices, which could additionally influence the performance of in-situ instruction techniques. Furthermore, we will switch from the relatively simple assembly task of placing LEGO DUPLO bricks to more complex task which, e.g., require bi-manual interaction and placements in more arbitrary orientations.

9 ACKNOWLEDGEMENTS

REFERENCES

[1] H. Alvarez, Iker Aguinaga, and Diego Borro. 2011. Providing guidance for maintenance operations using automatic markerless Augmented Reality system. In *2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 181–190. https://doi.org/10.1109/ISMAR.2011.6092385

[2] Alexander Bannat, Frank Wallhoff, Gerhard Rigoll, Florian Friesdorf, H. Bubb, Sonja Stork, H. J. Müller, Anna Schubö, Mathey Wiesbeck, and Michael F. Zäh. 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*. http://www.academia.edu/download/42003151/Towards_Optimal_Worker_Assistance_A_Fram20160203-32507-1ncxacj.pdf

[3] B. Besbes, S.N. Collette, M. Tamaazousti, S. Bourgeois, and V. Gay-Bellile. 2012. An interactive Augmented Reality system: A prototype for industrial maintenance training applications. In *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 269–270. https://doi.org/10.1109/ISMAR.2012.6402568

[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 (PETRA'17)*. ACM. https://doi.org/10.1145/3056540.3056547

[5] Markus Funk, Andreas Bächler, Liane Bächler, Oliver Korn, Christoph Krieger, Thomas Heidenreich, and Albrecht Schmidt. 2015. Comparing Projected In-situ Feedback at the Manual Assembly Workplace with Impaired Workers. In *Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '15)*. ACM, New York, NY, USA, 1:1–1:8. https://doi.org/10.1145/2769493.2769496

[6] Markus Funk, Thomas Kosch, Scott W. Greenwald, and Albrecht Schmidt. 2015. A Benchmark for Interactive Augmented Reality Instructions for Assembly Tasks. In *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia (MUM '15)*. ACM, New York, NY, USA, 253–257. https://doi.org/10.1145/2836041.2836067

[7] 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 (UbiComp '16)*. ACM, New York, NY, USA, 934–939. https://doi.org/10.1145/2971648.2971706

[8] D. Gorecky, M. Schmitt, M. Loskyll, and D. Zühlke. 2014. Human-machine-interaction in the industry 4.0 era. In *2014 12th IEEE International Conference on Industrial Informatics (INDIN)*. 289–294. https://doi.org/10.1109/INDIN.2014.6945523 00067.

[9] Anhong Guo, Shashank Raghu, Xuwen Xie, Saad Ismail, Xiaohui Luo, Joseph Simoneau, Scott Gilliland, Hannes Baumann, Caleb Southern, and Thad Starner. 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)*. ACM, New York, NY, USA, 71–78. https://doi.org/10.1145/2634317.2634321

[10] Ankit Gupta, Dieter Fox, Brian Curless, and Michael Cohen. 2012. Duplo-Track: 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 (UIST '12)*. ACM, New York, NY, USA, 389–402. https://doi.org/10.1145/2380116.2380167

[11] Julie Heiser, Doantam Phan, Maneesh Agrawala, Barbara Tversky, and Pat Hanrahan. 2004. Identification and Validation of Cognitive Design Principles for Automated Generation of Assembly Instructions. In *Proceedings of the Working Conference on Advanced Visual Interfaces (AVI '04)*. ACM, New York, NY, USA, 311–319. https://doi.org/10.1145/989863.989917

[12] S. Henderson and S. Feiner. 2011. Exploring the Benefits of Augmented Reality Documentation for Maintenance and Repair. *IEEE Transactions on Visualization and Computer Graphics* 17, 10 (2011), 1355–1368. https://doi.org/10.1109/TVCG.2010.245

[13] S. J. Henderson and S. Feiner. 2009. Evaluating the benefits of augmented reality for task localization in maintenance of an armored personnel carrier turret. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*. 135–144. https://doi.org/10.1109/ISMAR.2009.5336486

[14] B. M. Khuong, K. Kiyokawa, A. Miller, J. J. La Viola, T. Mashita, and H. Takemura. 2014. The effectiveness of an AR-based context-aware assembly support system in object assembly. In *2014 IEEE Virtual Reality (VR)*. 57–62. https://doi.org/10.1109/VR.2014.6802051

[15] Oliver Korn, Markus Funk, Stephan Abele, Thomas Hörz, and Albrecht Schmidt. 2014. Context-aware Assistive Systems at the Workplace: Analyzing the Effects of Projection and Gamification. , Article 38 (2014), 8 pages. https://doi.org/10.1145/2674396.2674406

[16] Oliver Korn, Markus Funk, and Albrecht Schmidt. 2015. Assistive systems for the workplace: Towards context-aware assistance. *Assistive Technologies for Physical and Cognitive Disabilities* (2015), 121–133.

[17] Oliver Korn, Albrecht Schmidt, and Thomas Hörz. 2013. Augmented Manufacturing: A Study with Impaired Persons on Assistive Systems Using In-situ Projection. In *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments (PETRA '13)*. ACM, New York, NY, USA, 21:1–21:8. https://doi.org/10.1145/2504335.2504356

[18] Robert E. Kraut, Susan R. Fussell, and Jane Siegel. 2003. Visual Information As a Conversational Resource in Collaborative Physical Tasks. *Hum.-Comput. Interact.* 18, 1 (June 2003), 13–49. https://doi.org/10.1207/S15327051HCI1812_2

[19] N. Petersen, A. Pagani, and D. Stricker. 2013. Real-time modeling and tracking manual workflows from first-person vision. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 117–124. https://doi.org/10.1109/ISMAR.2013.6671771

[20] Rupert Reif and Willibald A. Günthner. 2009. Pick-by-vision: augmented reality supported order picking. *The Visual Computer* 25, 5-7 (March 2009), 461–467. https://doi.org/10.1007/s00371-009-0348-y

[21] Leonardo Rodriguez, Fabian Quint, Dominic Gorecky, David Romero, and Héctor R. Siller. 2015. Developing a Mixed Reality Assistance System Based on Projection Mapping Technology for Manual Operations at Assembly Workstations. *Procedia Computer Science* 75 (Jan. 2015), 327–333. https://doi.org/10.1016/j.procs.2015.12.254

[22] Oliver Sand, Sebastian Büttner, Volker Paelke, and Carsten Röcker. 2016. smARt.Assembly – Projection-Based Augmented Reality for Supporting Assembly Workers. In *Virtual, Augmented and Mixed Reality (Lecture Notes in Computer Science)*, Stephanie Lackey and Randall Shumaker (Eds.). Springer International Publishing, 643–652. http://link.springer.com/chapter/10.1007/978-3-319-39907-2_61

[23] Angela L. Sauer, Andra Parks, and Patricia C. Heyn. 2010. Assistive technology effects on the employment outcomes for people with cognitive disabilities: A systematic review. *Disability and Rehabilitation. Assistive Technology* 5, 6 (2010), 377–391. https://doi.org/10.3109/17483101003746360

[24] Bjorn Schwerdtfeger and Gudrun Klinker. 2008. Supporting Order Picking with Augmented Reality. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '08)*. IEEE Computer Society, Washington, DC, USA, 91–94. https://doi.org/10.1109/ISMAR.2008.4637331

[25] D. Stanimirovic, N. Damasky, S. Webel, D. Koriath, A. Spillner, and D. Kurz. 2014. [Poster] A Mobile Augmented reality system to assist auto mechanics. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 305–306. https://doi.org/10.1109/ISMAR.2014.6948462

[26] Arthur Tang, Charles Owen, Frank Biocca, and Weimin Mou. 2003. Comparative Effectiveness of Augmented Reality in Object Assembly. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 73–80. https://doi.org/10.1145/642611.642626