ORIGINAL PAPER

Comparing head-mounted and handheld augmented reality for guided assembly

Nicolas Leins[1,2] · **Jana Gonnermann-Müller**[1,2] · **Malte Teichmann**[1,2]

Abstract

Different Augmented Reality (AR) displays are becoming more commonly used for work since AR promises benefits by offering support, e.g., with additional information or hints. However, most research compares AR with traditional work support, like paper-based or web-based instructions. Since various AR technologies offer device-specific advantages and disadvantages, different AR technologies are more or less suitable to offer support without overwhelming or distracting the worker. Research, therefore, needs to derive empirical results from comparing different AR displays to derive concrete recommendations for action on the use and design of AR for specific contexts. To address this research gap, this experimental study investigates the effect of video-see-through head-mounted AR (Varjo XR-3) vs. handheld AR (Apple iPad) on performance (time and committed failure), motivation, and cognitive load for guided assembly. The study results reveal that both AR displays can successfully guide people in guided assembly tasks. On a descriptive level, the head-mounted AR device reveals slightly better results in terms of time and committed failures. Notably, the impact of technical restrictions on the study results was still evident. Accordingly, further investigation of device-specific differences is of continuing importance.

Keywords Augmented reality · Assembly · Head-mounted display · Handheld display · Cognitive load

1 Introduction

New technologies and artificial intelligence have been introduced to the industry, making production more efficient and flexible [1]. The increasing variety and degree of individualization are also changing work environments, especially in industrial assembly and production. These changes force workers to learn new knowledge and skills to keep pace [2–4]. Augmented Reality (AR) can offer solutions for

✉ Nicolas Leins
nicolas.leins@wi.uni-potsdam.de

Jana Gonnermann-Müller
jana.gonnermann@wi.uni-potsdam.de

Malte Teichmann
malte.teichmann@wi.uni-potsdam.de

[1] Chair of Business Informatics, Processes and Systems, University of Potsdam, 14482 Potsdam, Germany

[2] Research Group Education for the Digital World, Weizenbaum Institute – Research for the Networked Society, 10623 Berlin, Germany

job-integrated support by presenting additional information and important cues needed in industrial workplaces [5]. Thereby, AR can support the worker, e.g., in the context of guided assembly [6]. However, traditional methods, like printed manuals or web-based information, are still commonly used to support the worker with additional information. AR offers the benefit of work-integrated information or visualizing abstract models in 3D, which is not possible with paper instruction or 2D displays. By overlaying or connecting the real world with virtual objects, instructions can be displayed directly in the user's field of view (FOV) and related to the real environment [7]. Wu et al. summarized that AR technologies "enable (1) learning content in 3D perspectives, (2) ubiquitous, collaborative and situated learning, (3) learners' sense of presence, immediacy, and immersion, (4) visualizing the invisible, and (5) bridging formal and informal learning" [8].

Looking at the benefits of using AR, recent meta-analyses confirm the effect of AR over traditional learning methods in various contexts [5, 9, 10]. However, besides the potential benefits of using AR on performance, research has shown inconsistencies concerning the use and design

of AR for practical use. While study results show that the additional information provided by AR fosters performance and reduces the cognitive load on workers, other studies show that the additional information can also provoke overload. Research claims that although in-place information, image recognition, and gaming elements offer added value over traditional methods, they can also distract and overwhelm the worker [11]. In addition to the global question of whether AR adds value, further questions arise regarding which type of AR device (head-mounted or handheld displays, spatial AR screens, or projections) is suitable for which context and how AR must be designed [12]. Howard and Davis [5] encourage testing the impact of specific features on cognitive load to test recommendations for the use and design of AR in terms of effectiveness and performance. In addition to design principles and the question of how to use them meaningfully, it is initially the characteristics of the technologies that create possibilities or limit AR in certain situations, for example, through bad resolution or a small FOV.

Therefore, the basis of any further use is in the experimental testing of technologies and their possibilities [13]. This experimental study compares two types of AR displays, head-mounted (HMD) and handheld AR (HHD), on performance (time and failures) in a guided assembly task. To contribute to the discourse on inconsistencies regarding AR's impact on cognitive load, this study experimentally investigates whether HMD or HHD differently affect cognitive load, which constitutes an influential aspect on the side of the user that influences AR technology's effective use and acceptance [14].

1.1 Augmented reality in guided assembly

Guided assembly is defined as the process whereby an individual assembles parts or components into a final product, following detailed instructions or visual aids. This approach is frequently employed in manufacturing or construction to guarantee accuracy and consistency. The instructions may be delivered through manuals, digital interfaces, or augmented reality systems. Research shows that AR instructions improve performance (e.g., assembly time and failures) of guided assembly tasks compared to traditional information support [15–20]. Studies confirmed that AR reduces the cognitive load during assembly [16, 17, 21]. At the same time, other research also shows increased cognitive load caused by using AR in the work process [22, 23]. A representative survey conducted on AR shows doubts on the worker side about the regular use of AR due to the expectation of overload, distraction, and disruption [24]. Concerns about the use can be made up of various factors. In addition to the perceived added value that AR brings to performance,

it is also important for people how easy or challenging it is to use AR regularly. The cognitive load indicates the cognitive demand of a specific task or information (system) on the cognitive system [25]. In education and learning research, the cognitive load theory describes the cognitive demand on the learner. The theory is based on assumptions about the human cognitive architecture and the restricted capacity of the working memory that allows only a limited amount of information to be processed at a specific time [26]. A similar conceptualization of cognitive load is coined by Paas et al. [27], who differ between mental effort, which describes cognitive load linked to learner aspects, and mental demand linked to the task. Cognitive load is commonly measured using questionnaires or physiological assessment methods, such as eye-tracking, skin conductance, or EEG [28, 29]. The underlying assumption is that by reducing unnecessary information or additional distractions, the cognitive demand on the worker can be reduced, which leaves enough capacity for the task.

AR devices are distinct in terms of their type of display, e.g., HMD, HHD, and spatial displays [30]. HMDs are worn on the head, allowing virtual content to be displayed directly in the user's FOV while the hands remain free for other tasks. Modern devices have binocular optics, which enable stereoscopic vision whereby the real and virtual worlds can be combined correctly in terms of perspective. In HMD, the merging of the real and virtual worlds is realized with the optical-see-through (OST) or video-see-through (VST) technique. The OST technique displays the virtual objects in the user's field of view using a transparent display. An optical overlay achieves the combination of the virtual and real environment. In VST, the real environment surrounding the user is captured by cameras and digitally combined with the virtual environment. The image is then displayed to the user via conventional displays (e.g., Fig. 5). HHDs are mobile devices that can be held by the user (usually using smartphones or tablets). HHD devices also use the VST technique to achieve augmentation through cameras attached to the back of the devices (e.g., Fig. 3). Compared to that, spatial Augmented Reality (SAR) technologies are detached from the user and integrated into the real environment. They also use the VST and OST as well as projection-based techniques. VST SAR devices are, for example, stationary computer screens that resemble HHD devices that cannot be moved around. OST SAR follows the same principle but uses technologies like transparent screens or holograms. The last category is projection-based techniques, which project images directly onto the surface of physical objects. Figure 1 gives an overview of AR display classifications.

When examining guided assembly tasks, most studies focus their consideration of AR technologies on the
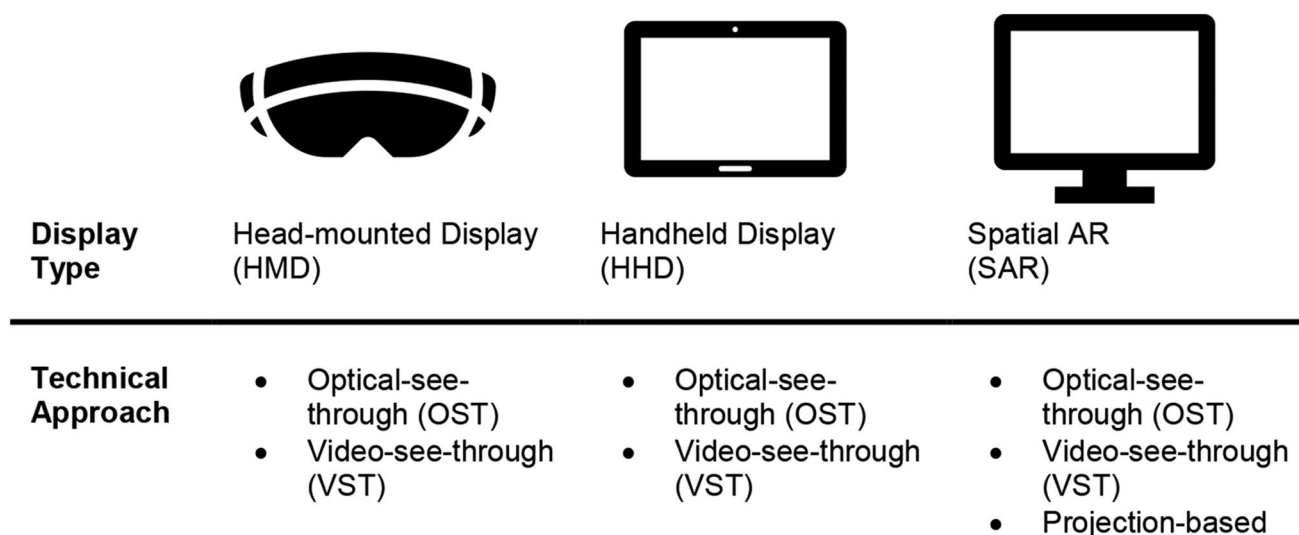
Display Type	Head-mounted Display (HMD)	Handheld Display (HHD)	Spatial AR (SAR)
Technical Approach	• Optical-see-through (OST) • Video-see-through (VST)	• Optical-see-through (OST) • Video-see-through (VST)	• Optical-see-through (OST) • Video-see-through (VST) • Projection-based

Fig. 1 Classification of augmented reality displays (based on Carmigniani et al. [30])

comparison to conventional transfer instruments (e.g., paper instructions, electronic instructions on computer/tablet). The comparison of AR devices among each other is conducted less extensively [31]. Since AR can be seen more as a concept than a specific technology [8], the different AR hardware solutions differ significantly. This diversity of AR devices must be considered when using them as support in guided assembly tasks. Besides the differences in displays (HHD, HMD, and SAR), the degree of immersion and features of the specific AR hardware differ (e.g., in terms of display resolution, camera quality, or FOV).

Given the diversity of different AR devices, it is insufficient to use studies that examine differences from traditional forms of information delivery only at the level of global technologies (e.g., AR vs. web-based or paper-based instructions). Rather, it seems important to break down the complexity into its individual parts and investigate the individual AR display experimentally to make statements about which aspects of the technology lead to improving performance and minimizing overload or, conversely, contribute to overload. It is only through a detailed analysis that it becomes clear which technology is suitable for which application.

1.2 Effects of augmented reality

Studies have already investigated different AR aspects in the context of work and assembly. The following section summarizes the findings of a literature review of studies investigating differences between AR displays on performance, which builds the foundation for the study presented.

Alves et al. [32] investigated HHD AR and projection-based SAR assisting a puzzle assembly. The group using SAR needed significantly shorter assembly time and committed fewer failures. The authors also investigated the demand during the task by measuring the cognitive load (NASA Task Load Index (TLX)). Participants reported a slightly lower cognitive load with SAR.

Alves et al. [33] compared HHD AR, OST HMD AR, and VST SAR (stationary monitor) in a LEGO assembly task. SAR and HMD AR led to significantly shorter assembly times than HHD AR. Regarding the cognitive load (Raw TLX), participants using SAR reported significantly less cognitive load.

Blattgerste et al. [34] compared conventional paper instructions with HHD AR using a smartphone and two OST HMDs using a Microsoft HoloLens and an Epson Moverio BT. People using the paper instructions showed the fastest assembly time, while people in the OST HMD group with the HoloLens made the fewest failures. Cognitive load was significantly lower in the paper instructions group compared with all the AR instruction groups. People using the Epson Moverio BT-200 reported significantly lower cognitive load (Nasa-TLX) than those with HHD smartphones.

In a study by Büttner et al. [35], projection-based SAR was compared with OST HMD AR using a Vuzix STAR 1200. The OST HMD glasses produced worse results than projection-based SAR in a Lego assembly task.

This literature overview depicts differences between AR displays on performance and cognitive load. However, only devices with OST technology were used for HMD AR, which shows restrictions of small FOVs OST HMD AR devices. The FOV of the Vuzix STAR 1200 (35° diagonal FOV) used by Büttner et al. [35] and the Microsoft HoloLens (35° diagonal FOV) and Epson Moverio BT (23° diagonal FOV) used by Blattgerste et al. [34] cover only a fraction of the human FOV. The studies emphasize the small FOV of the HMD AR devices as a technical limitation, which may

be one of the reasons for not performing as well as other AR technologies [34, 35]. This is of major importance since research shows the influence of the size of the FOV in HMD VR [36] and HMD AR [37–39]. After summarizing the discourse on AR, we emphasize the following three gaps we address with our study:

- Different AR displays are rarely compared experimentally between each other.
- Results of comparative experiments are highly dependent on the used devices and investigated setting.
- Investigated HMD AR devices often reveal their narrow FOV as a performance-limiting factor.

1.3 Aim of the study

Building on the identified research gaps, this study conducts a comparative analysis of the effects of two distinct AR displays (VST HMD vs. VST HHD) on performance in guided assembly by conducting a controlled experiment. Unlike previous research [15–20] that primarily compares AR with other media (e.g., paper, tablets, computers), this study addresses a critical limitation: conclusions drawn from such comparisons tend to generalize AR as a whole without offering insights into specific AR features. This approach overlooks important distinctions, such as the strengths and weaknesses of different display types or see-through technologies. The AR devices that we investigate in this study are distinguished by two key characteristics. Firstly, the use of a HMD device allows for hands-free interaction, whereas a HHD device necessitates the use of hands. Secondly, the employed HMD device incorporates high-quality VST technology, a relatively under-researched area. The primary advantage of this technology lies in its ability to provide a wider field of view (FOV), which is often regarded as a limiting factor for HMD devices. By employing an experiment-based methodology, this study enables a more precise examination of two different AR displays, facilitating conclusions regarding their specific advantages and disadvantages within the targeted use case. On this basis, the following three research questions are addressed:

RQ1: Do HHD and HMD AR devices differ in their impact on assembly performance (assembly time and failures)?

In order to enhance assembly performance through the use of AR guidance, it is important to consider the cognitive load of the user [5]. To improve user experience, an effective AR system should avoid distractions or overwhelming the user. Therefore, the cognitive load is measured to investigate the second research question:

RQ2: Do HHD and HMD AR devices evoke differences in the user's cognitive load depending on the AR technology used?

In addition to cognitive load, motivation is also considered as a determining success factor in multimedia environments. Multimedia environments offer information in at least two modes: text, picture, audio, or haptics [40]. Based on that, cognitive load and motivation should be considered in the instructional design of multimedia environments such as AR [41]. Research indicates that motivational factors act as a mediator of learning by increasing or decreasing cognitive engagement [42]. This effect could be beneficially influenced by the type of AR device used. Therefore, the third research question investigates participants' motivation after using the AR technologies to cover both aspects:

RQ3: Do HHD and HMD AR devices affect the user's motivation using AR?

To address the research questions, the present study adopts the research design of the study by Grum and Gronau [43], "Adaptable Knowledge-Driven Information Systems Improving Knowledge Transfers", in which an HHD AR instruction for assembling a cupboard has been developed. The study demonstrated the fundamental usability of the HHD AR instruction by comparing it with conventional paper instruction. By building on the validated study design, this paper broadens the study's scope by comparing VST HMD AR with HHD AR. For this purpose, the Varjo XR-3 was chosen as a VST HMD AR device. The Varjo XR-3 has a large FOV of 115° horizontally and 90° vertically (approximately 146° diagonally), which is significantly larger than previously investigated HMD AR devices.

2 Method

To address the research questions, an experiment was conducted using a between-subjects design. Participants were randomly assigned to one of the two experiment groups where the AR device (HHD, HMD) was manipulated as an independent variable. The two experimental groups are abbreviated as HHD (n = 16) and HMD (n = 17) in the following. HHD stands for the group that used the handheld-display device (Apple iPad 10.2") for the assembly, and HMD stands for the group that used the head-mounted-display device (Varjo XR-3). As dependent variables, assembly time, failures, cognitive load, and motivation were measured. For this study, the experiment design of Grum and Gronau [43] was replicated, which was successfully validated in previous experiments. Grum and Gronau [43]

investigated the guided process of assembling the METOD cupboard from the company IKEA. In the following sections, the experiment design, the measured variables, and data analysis are described in more detail.

2.1 Participants

The investigated sample ($N=33$) consisted of 15 women (45.5%) and 18 men (54.5%) with an average age of 25.6 years (women: 26 years; men: 25.2 years). Participants were employees and students of the University of Potsdam, recruited via a mailing list and flyers. Under randomized assignment, the participants were divided into control and experimental groups according to the two-group study design. Within the experimental groups, there were 11 women (69%) and five men (31%) in the HHD condition and four women (24%), and 13 men (76%) in the HMD condition.

2.2 AR instruction

The AR instruction for assembling the cupboard was adopted from the study by Grum and Gronau [43]. The instruction was based on the original paper instruction for the cupboards provided by the company IKEA and developed for an HHD AR device. For this study, the AR instruction had to be transferred to a VST HMD AR Device (Varjo XR-3). The two AR applications are explained in more detail in the following.

2.2.1 HHD AR application

The previously validated HHD AR application of the study by Grum and Gronau [43] is reused for the study without any modifications. It is available under the GNU Affero General Public Open Source License [44]. The AR application is a step-by-step instruction for assembling a METOD cupboard from the company IKEA. Figure 2 depicts the eight instruction steps.

The application was developed with Unity and uses the Vuforia Software Development Kit to develop AR applications. The application visualizes the instruction steps with 3D models of the cupboard components and the tools needed for assembly. Using buttons in the upper left and lower right corners, users can navigate through the individual steps of the instructions. In the process, the corresponding 3D models are augmented directly onto the real components of the cupboard. This uses image targets, which have been placed on the cupboard components so that the 3D models match the real components precisely in size and position. The instruction shows in each step which components must be attached at which position. In addition, instructional text for the specific step was displayed at the top of the screen. The application was installed on an Apple iPad 10.2" for the experiment. Figure 3 shows the exemplary usage of the HHD AR application.

2.2.2 HMD AR application

The instructional content is adapted for use on the HMD device. The quality and quantity of the instructional content remained the same, ensuring no differences between HHD and HMD conditions. Due to technology-specific
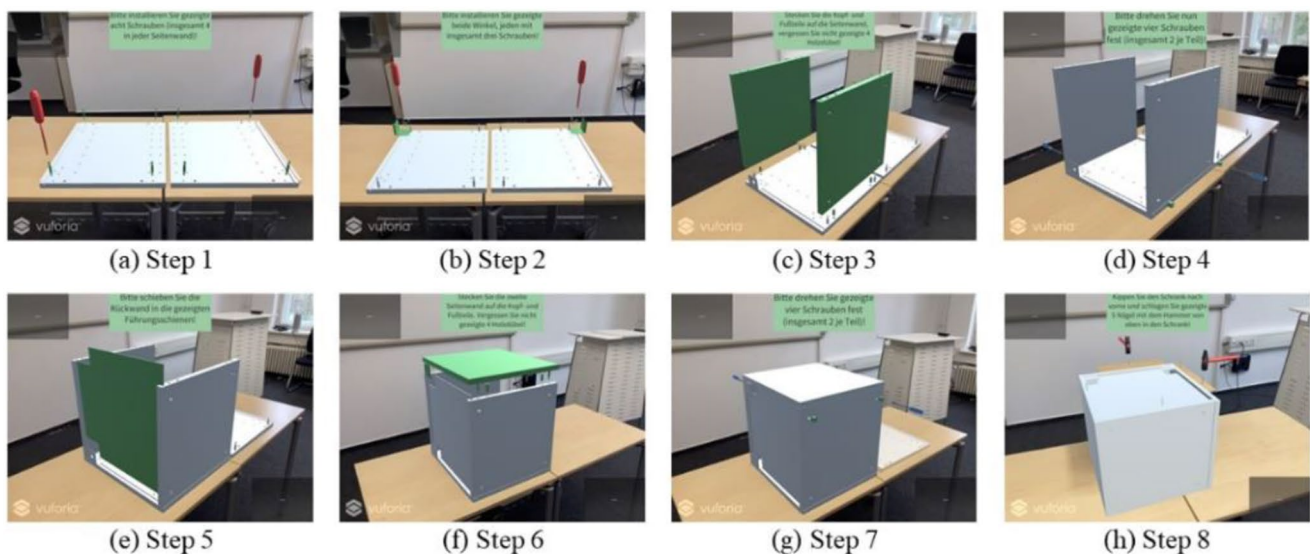

Fig. 2 Overview of the AR instruction for the different assembly steps

(a) Step 1   (b) Step 2   (c) Step 3   (d) Step 4

(e) Step 5   (f) Step 6   (g) Step 7   (h) Step 8
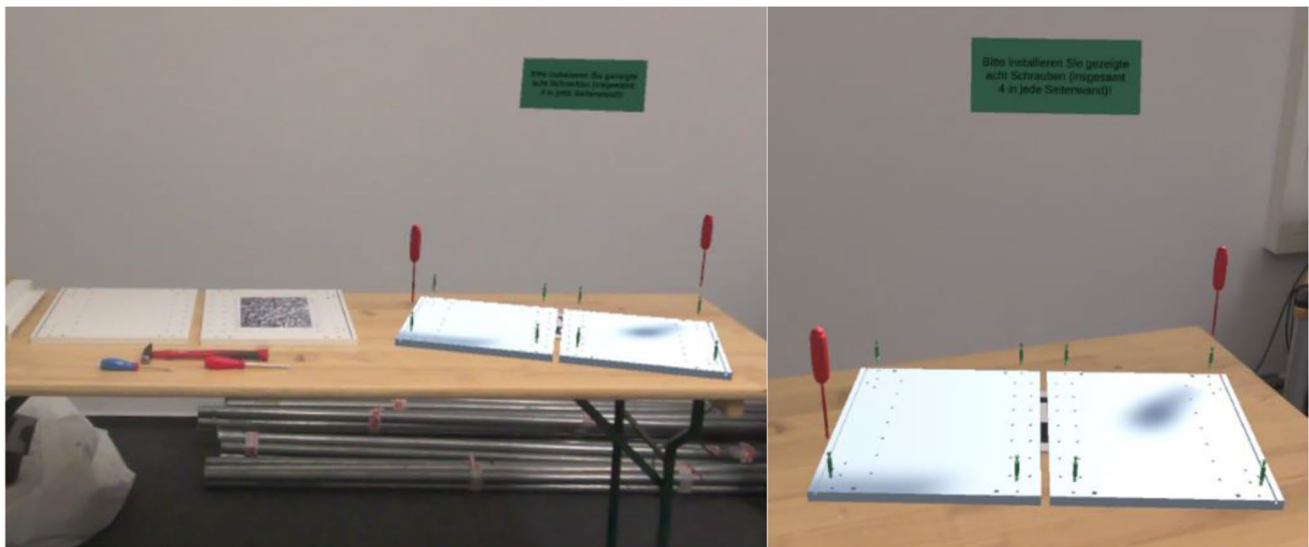
Fig. 3 HD AR application on Apple iPad 10.2"

Fig. 4 Augmentation of the assembly instruction next to the real assembly objects (HMD AR application)

differences between the two AR devices, the following three adjustments had to be made to keep the two conditions constant. In the following, the three adjustments are explained in more detail:

(1) To avoid obscuring the real components through the augmented objects and, therefore, to enable wearing the HMD AR device continuously during the assembly process, the instructional steps were not augmented directly onto the real components of the cupboard. For these reasons, it was decided to augment the instruction steps directly next to the actual components (see Fig. 4). This made wearing the HMD AR device possible throughout the assembly process (Fig. 5). Putting the AR glasses aside during the assembly process, as could be done with the HHD tablet, should be

avoided to rule out any interfering factors that might occur in the process.

(2) The implementation of a controller system, which enables participants to navigate through the instruction steps, was omitted in the HMD AR application. Performing a pre-test, two control options (hand tracking and voice control) were tested before. Since both control options were assessed as not intuitive and not user-friendly, they were not used in the main experiment. Both pre-tested control options provoked an increase in cognitive load, which should be avoided as a potential co-factor influencing the study results in an uncontrolled way. For this reason, the experimenter, who was present in the same room, navigated the HMD AR application. Following the participant's completion of an assembly step, the experimenter was informed

Fig. 5 Participant wearing the HMD AR glasses (Varjo XR-3) while assembling the cupboard

and proceeded with the instruction to the next step from a remotely connected computer. In the HHD AR application, the participants navigated over the intuitive touch buttons, which did not cause any problems in the previously conducted experiment by Grum and Gronau [43]. In both experimental groups, participants were required to provide information regarding a completed instructional step due to the time measurement (further described in Sect. 2.3). This ensured that the two experimental groups were comparable.

(3) In the eighth instruction step, the participants were asked to tilt the cupboard forward to hammer the nails into the back wall. Due to the attached image target on the back of the cupboard, the HHD application tilts the instructions and displays them correctly in perspective on the cupboard. Since the instructions are displayed independently of the actual cupboard in the HMD AR application, it was decided to visualize the tilting of the cupboard by animating the 3D model. The animation started once at the beginning of the instruction step and rotated the model of the cupboard forward by 90°.

The HMD application was also developed using Unity with the provided Varjo plugin. An object marker from Varjo intended for this purpose was used to place the 3D models of the instructions next to the real cupboard components. This allowed the participants to stand in front of the cupboard components while seeing the instructions by looking to the right side (see Fig. 4). The instruction texts of each step were adopted to ensure similar instruction between HMD and HHD. They were displayed on a text panel above

the 3D models of the instructions (see Fig. 4). For the navigation, a control option over the keyboard of the connected computer was implemented to enable the experimenter to navigate through the instruction. Additionally, SteamVR with two base stations 2.0 from HTC was used to improve tracking with the Varjo XR-3.

2.3 Procedure

At the beginning of the experiment, the participants were informed about the experiment and agreed to voluntary participation and specifics associated with data acquisition. Afterwards, the study's overall intent and experimental procedure were explained to the participants. Before the start of the experiment, the HMD AR device was attached to the head and calibrated to ensure a correct operation. In addition, participants were informed about the control of the instruction and had time to familiarize themselves with the device and ask questions about the experimental procedure. Regarding the assembly process, the participants were advised to follow the instructions exactly. Once all ambiguities were cleared, the experiment started. Figure 6 shows the initial positioning of the required cupboard components, which was the same for all study participants. Depending on the group, the step-by-step assembly of the cupboard was performed either with the help of the HHD or HMD AR instruction. During assembly, the experimenter observed the process and documented the failures and time the participants required for each instruction step. With the help of a
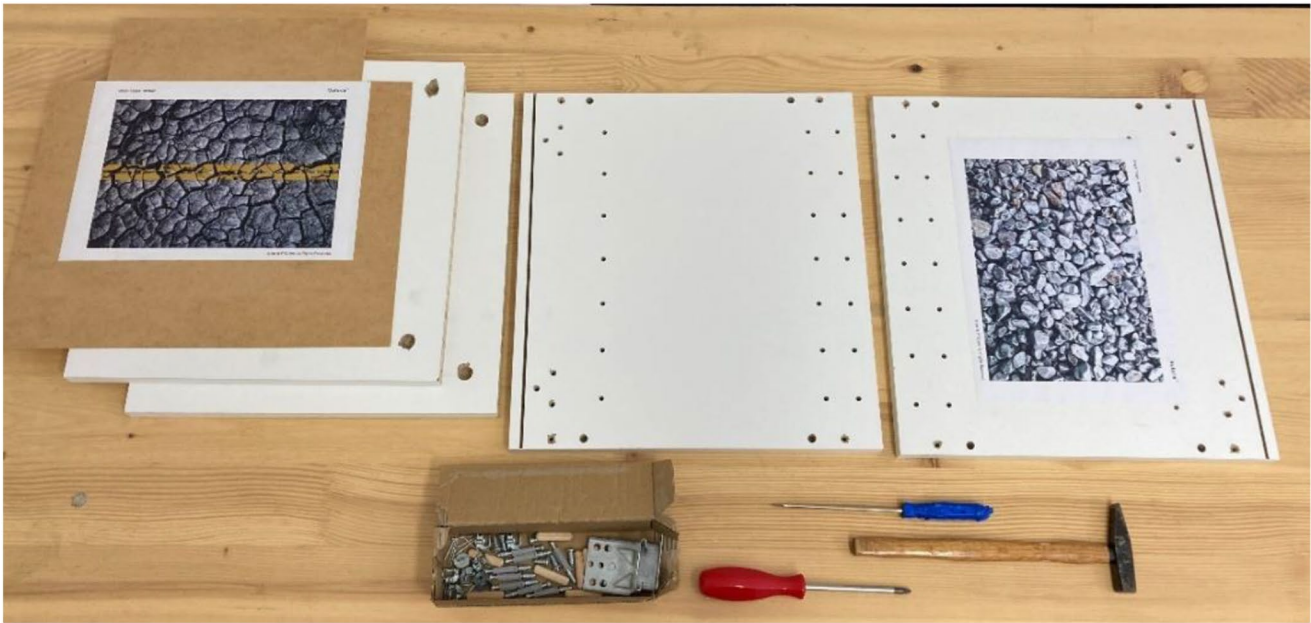
Fig. 6 Initial positioning of the cupboard components

self-developed form, the experimenter filled out a standardized documentation during the experiment (see Sect. 2.4). At the end of the experiment, each participant filled out the final questionnaire after the cupboard had been fully assembled. To ensure consistency across all participants, the greeting, briefing, explanation of the experimental procedure, and calibration were conducted following a standardized protocol. This process and the documentation of errors and time were systematically recorded in an experimental log to maintain uniform conditions throughout the study.

2.4 Measurements

Building on Grum and Gronau [43], an experimental log was developed to systematically document time and failures during the assembly process in a standardized manner. The experimental log was used to separately document the failures committed and the time required for every instruction step. The log lists typical failures for each assembly step identified during the conducted pre-test. If a participant made one of the listed failures, the experimenter could mark the failure as occurring in the corresponding column. If the same failure was made multiple times in the same assembly step, it was also possible to indicate how often it occurred. For example, in the third assembly step, two side panels must be assembled. If one of the two side panels was mounted incorrectly, it was counted as one failure. On the other hand, two failures were documented if both side panels were mounted incorrectly. Rare failures not listed in the questionnaire were added using free text fields.

Table 1 Minor and major failures in the assembly [43]

Assembly step	Minor failure	Major failures
1	Wrong tool used Forgot to mount screw	Screw mounted in wrong position Dowel instead of screw Nail instead of screw
2	Wrong tool used Forgot to mount screw	Angle incorrectly mounted Angle mounted in wrong position Forgot angle
3	Wrong tool used Forgot to mount dowel	Forgot to mount side wall Wrong position of side wall Side wall wrong way around
4	Wrong tool used Forgot to mount screw Forgot to tighten screw	-
5	-	Forgot back panel Back panel not properly mounted Back panel wrong way around
6	Forgot to mount dowel	Side wall wrong way around Dowel in wrong position
7	Wrong tool used Forgot to mount screw Forgot to tighten screw	-
8	Wrong tool used Nails in wrong position Too many nails Too few nails	-

Following Grum and Gronau [43], the failures were divided into major and minor failures (listed in Table 1). Minor failures do not affect subsequent assembly steps, while major failures prevent subsequent steps. In case a

major failure occurred, the experimenter informed the participants at the end of an assembly step. To proceed, the mistake had to be corrected by the participant before proceeding with the next step. The time required for correction was added to the time previously needed. The experimenter documented minor errors that did not have to be corrected.

Time measurement was also conducted separately for each step. Before the start of the experiment, the study participants were instructed to signal the experimenter when an assembly step was to be started and when it was estimated to be completed. The experimenter used this as a start and stop signal for manually recording time using a stopwatch. At the end of each assembly step, the time was rounded to whole seconds and documented in the questionnaire. After the experiment, participants completed a questionnaire to assess their cognitive load and motivation during the assembly process.

The single-item scale from Paas was used to measure the self-assessed cognitive load in the post-questionnaire [45]. A 9-point Likert scale was used to assess how high the mental effort was during the task ("Very, very low mental

Table 2 Intrinsic motivation inventory (IMI) scales and questions used [46]

IMI scale	Questions
Interest/enjoyment	I enjoyed doing this activity very much.
	This activity was fun to do.
	I thought this was a boring activity. (R)
	This activity did not hold my attention at all. (R)
	I would describe this activity as very interesting.
	I thought this activity was quite enjoyable.
	While I was doing this activity, I was thinking about how much I enjoyed it.
Perceived competence	I think I am pretty good at this activity.
	I think I did pretty well at this activity, compared to other students.
	After working at this activity for awhile, I felt pretty competent.
	I am satisfied with my performance at this task.
	I was pretty skilled at this activity.
	This was an activity that I couldn't do very well. (R)
Effort/importance	I put a lot of effort into this.
	I didn't try very hard to do well at this activity. (R)
	I tried very hard on this activity.
	It was important to me to do well at this task.
	I didn't put much energy into this. (R)
Pressure/tension	I did not feel nervous at all while doing this. (R)
	I felt very tense while doing this activity.
	I was very relaxed in doing these. (R)
	I was anxious while working on this task.
	I felt pressured while doing these.

(R) indicates that the item has to be reversed

effort." (1) to "Very, very high mental effort." (9)). Since the study was conducted in German, the question and answer options were translated into German.

A questionnaire from the Intrinsic Motivation Inventory [46] was created to measure the participants' motivation. The Intrinsic Motivation Inventory (IMI) is a multidimensional set of scales and questions that can be used to measure participants' subjective experience during an activity in laboratory experiments. Since the questionnaire allows for selecting subscales tailored to the respective application context, the subscales *Interest/Enjoyment*, *Perceived Competence*, *Effort/Importance*, and *Pressure/Tension* were used for this study. The scales and the corresponding questions are listed in Table 2. The *Interest/Enjoyment* scale directly captures the intrinsic motivation. In addition, the *Perceived Competence* scale was used, which is considered a positive predictor of intrinsic motivation (*Interest/Enjoyment* scale). On the other hand, the questions from the *Pressure/Tension* scale represent a negative predictive value. The *Effort/Importance* scale must be considered separately. It is used to record an expected consequence of intrinsic motivation. As required by the IMI, the 23 questions were arranged in random order and had to be answered using a 7-point Likert scale. The participants were asked to assess the extent to which the statements applied to them in consideration of the assembly task they had completed. The answer options ranged from "not true at all" (1) to "absolutely true" (7). Since the study was conducted in German, the IMI questionnaire translated into German was used [46].

2.5 Data analysis

The statistical analysis of the collected data was performed using the R programming language. The different data sets were first linked through the individual code during the data cleaning process. For the failure measurement, the sum of all minor and major failures was calculated for each step. The questions of the IMI question catalog also had to be preprocessed for the analysis. First, the ratings of the corresponding questions had to be reversed (see Table 2). Then, the scale rating could be calculated for each of the four subscales using the mean value of all questions on the respective scale. Only the scale ratings of the subscales *Interest/Enjoyment*, *Effort/Importance*, *Perceived Competence*, and *Pressure/Tension* were used for further analysis.

In addition to descriptive statistics of the sample and the results of the time, failures, motivation (IMI scales), and cognitive load (Paas' scale) measurements, significance tests were calculated to test for differences between the groups. In advance, all requirements for running the intended statistical calculations were checked. Due to the normally distributed time variable, a two-sided Welch's t-test was applied to

Table 3 Time measurement per assembly step, descriptive statistics, and comparison of means with Welch's t-test

Assembly step	Mean time (seconds)		Standard deviation		t-test (two-sided)		
	HHD $n=16$	HMD $n=17$	HHD $n=16$	HMD $n=17$	t	Degrees of freedom	p-value
Step 1	124	134	22.47	27.85	-1.07	30.33	0.294
Step 2	137	128	25.24	26.26	0.94	30.98	0.353
Step 3	90	97	29.67	34.14	−0.57	30.82	0.574
Step 4	76	70	35.88	15.77	0.57	20.30	0.573
Step 5	42	35	18.31	15.52	1.11	29.48	0.275
Step 6	59	49	18.32	14.27	1.69	28.32	0.102
Step 7	57	50	14.22	11.52	1.59	28.88	0.122
Step 8	106	89	33.05	20.59	1.68	24.82	0.105
Total	692	652	140.22	121.61	0.83	29.75	0.413

*** $p<.001$, ** $p<.01$, * $p<.05$
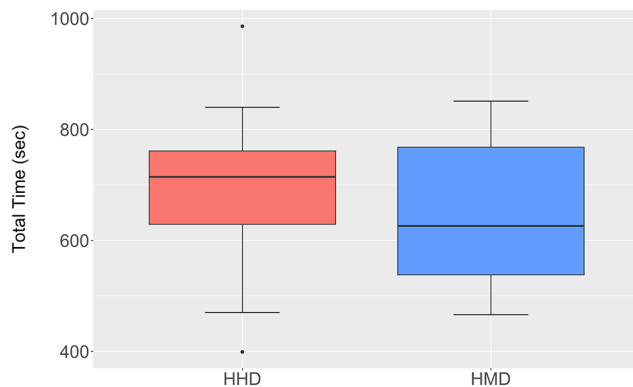
Fig. 7 Total assembly time (in seconds) for the HHD group (red) and the HMD group (blue)

Fig. 8 Assembly time (in seconds) per assembly step for the HHD and HMD condition (with error bars: ± SD)

test for differences between the groups (HHD and HMD). As the normal distribution assumption was not met for the other variables (failures, IMI scale ratings, and cognitive load), a two-sided Mann-Whitney-U-test was applied to test for group differences.

Additionally, Kendall's rank correlation coefficient [47] was used to provide an overview of the relationships between the two factors, motivation, and cognitive load, with performance (time and failure) in the assembly process. The significance level was set at 5%.

3 Results

To answer the research questions, the results regarding time and failures (RQ1), cognitive load (RQ2), and motivation (RQ3) are presented descriptively and compared inferentially between the two groups (HHD and HMD).

3.1 Assembly time

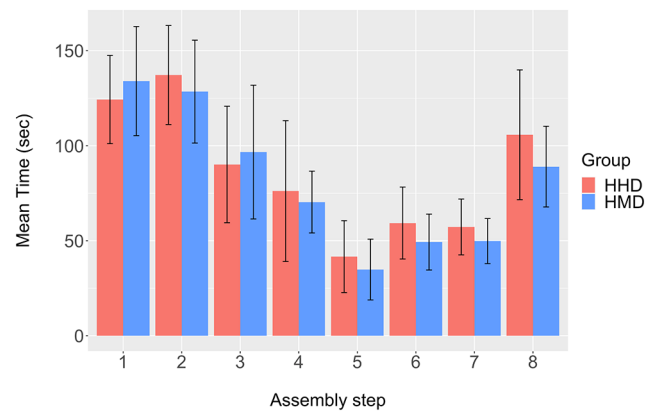Table 3 shows the assembly time for the groups HHD and HMD in seconds. The HHD group (M=692, SD=140.22)

required more time than the HMD group (M=652, SD=121.61) regarding the entire assembly process. Figure 7 visualizes the time difference for the assembly process between HHD and HMD. Considering each assembly step separately, steps 1 and 3 differ from the general effect. For steps 1 and 3, the HHD group required less time than the HMD group. For all other steps, the time advantage was with the HMD group. Figure 8 shows the time comparison between both groups for each assembly step.

The difference between HHD and HMD on the level of the entire assembly process and the individual assembly steps did not reach statistical significance (Table 3 shows the results of the two-sided Welch's t-test with t-value, degrees of freedom, and significance (p-value)).

3.2 Failures

The results of the failure measurement are listed in Table 4 and visualized in Fig. 9. On a descriptive level, considering the entire assembly process, the participants in the HMD group (M=0.94, SD=1.39) committed fewer errors than participants in the HHD condition (M=1.75, SD=1.88).

Table 4 Failure measurement per assembly step, mean values and Mann-Whitney-U-test

Assembly step	Mean failures		Mann-Whitney-U-test (two-sided)			
	HHD $n = 16$	HMD $n = 17$	U ($N_{HHD} = 16$, $N_{HMD} = 17$)	z	p-value	r
Step 1	0.13	0.00	144.5	−0.97	0.332	0.169
Step 2	0.06	0.00	144.5	−0.97	0.332	0.169
Step 3	0.00	0.41	104	-2.0	0.045*	0.348
Step 4	0.75	0.09	190	-2.29	0.022*	0.399
Step 5	0.13	0.00	153	-1.44	0.151	0.250
Step 6	0.00	0.00	136	-	1	-
Step 7	0.38	0.12	171	-1.67	0.094	0.291
Step 8	0.31	0.24	145.5	−0.46	0.649	0.079
Total	0.75	0.94	175.5	-1.47	0.14	0.257

*** $p < .001$, ** $p < .01$, * $p < .05$
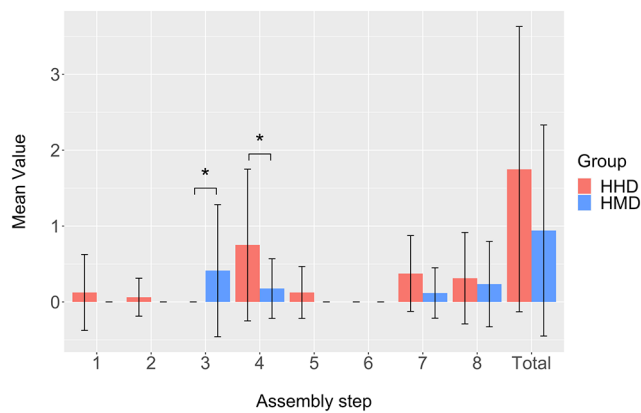
Fig. 9 Mean values of committed errors per assembly step for groups (HHD and HMD). *** $p < .001$, ** $p < .01$, * $p < .05$. Error bars: ± SD

Fig. 10 Self-assessed cognitive load for groups HHD and HMD ranging from 1 (low) to 9 (high)

Regarding the individual assembly steps, steps three and four stand out with the highest difference between HHD and HMD. While in step 3, none of the participants using the HHD committed a failure (M = 0), nearly every second made a mistake in the HMD condition. In contrast, in assembly step 4, almost none of the participants using HMD made a mistake (M = 0.09), while the HHD group committed failures (M = 0.75). The number of errors significantly differed between the groups for both assembly steps. Using HHD outperforms HMD in assembly step 3 ($p < .05$) while using HMD outperforms HHD in assembly step 4 ($p < .01$).

3.3 Cognitive load

Both groups report a relatively low cognitive load after the guided assembly process (HHD group M = 2.44; HMD group M = 2.71). The descriptive results are shown in Fig. 10. The results of the calculated Mann-Whitney-U-test did not suggest a statistically significant difference between HMD and HHD (U ($N_{HHD} = 16$, $N_{HMD} = 17$) = 121.5, z = −0.53, $p > .05$). Regarding a possible relationship between cognitive load and performance, the results showed no significant
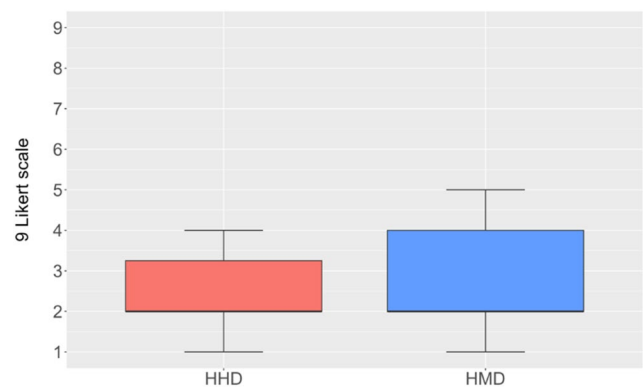
association neither between cognitive load and assembly time ($\tau c = 0.14$, $p > .05$) nor cognitive load and committed failures ($\tau c = 0.20$, $p > .05$).

3.4 Intrinsic motivation inventory (IMI)

The following section reports the results of the IMI subscales *Interest/Enjoyment*, *Effort/Importance*, *Perceived Competence*, and *Pressure/Tension*. The scale ratings range from 1 - low to 7 - high. For both groups, participants reported high ratings on the *Interest/Enjoyment* scale ($Md_{HHD} = 6.00$, $Md_{HMD} = 5.86$) and the *Perceived Competence* scale ($Md_{HHD} = 5.50$, $Md_{HMD} = 5.50$). The *Effort/Importance* scale showed slightly lower ratings, with $Md_{HHD} = 4.10$ and $Md_{HMD} = 4.80$. Ratings on the scale *Pressure/Tension* was lowest with Md = 2.20 for both conditions. The differences between HHD and HMD regarding the scale ratings were only marginal, and statistical tests did not reveal a significant difference.

Testing the scale ratings on correlation with performance (time and failures), the *Pressure/Tension* scale showed a significant relationship with assembly time ($\tau c = 0.33$, $p < .01$), indicating that longer assembly times are positively

associated with higher ratings in the perceived *Pressure/Tension* scale. All other scales showed no significant correlations, either on time or failures.

4 Discussion

This study aims to compare handheld display (HHD) and head-mounted display (HMD) AR devices to gain insights into the appropriateness of different AR devices for guided assembly. Therefore, this study included a VST HMD AR device (Varjo XR-3), which has rarely been investigated. The VST HMD device was compared with a handheld device to understand different technology capabilities and the limitations of different devices (field of view, resolution, interaction, haptics). To experimentally compare the devices in the application context of a guided assembly, the study builds on the work of Grum and Gronau [43], where an AR instruction for a HHD AR device was validated. After transferring the instruction to a VST HMD device, the present study addressed three research questions - *Do HHD and HMD AR devices differ in their impact on assembly performance (assembly time and failures)?; Do HHD and HMD AR devices evoke differences in the user's cognitive load depending on the AR technology used?* and *Do HHD and HMD AR devices affect the motivation of the user's motivation using AR?*

Performing the time and error measurements again with the HHD AR guidance yielded similar results to those obtained by Grum and Gronau [43]. The study results also highlight that the instruction could be successfully transferred between the previously used HHD AR and the VST HMD AR device. Results regarding the investigated research questions are discussed in more detail in the following sections, and the research questions are answered.

4.1 Impact of AR devices on assembly time and failures

Assembly time and failures were measured to assess performance using two different AR devices for the guided assembly process. Similar findings for both groups indicate that both devices can be successfully used for the applied guided assembly scenario. The participants needed, on average, 692 s (HHD) and 652 s (HMD) to complete the entire assembly process. No statistically significant results could be found between the groups. Looking into the failures, the participants committed 1.75 (HHD) and 0.94 (HMD) failures on average, which can be considered as relatively few failures, but also means that on average, none of the participants has assembled the cupboard without failures. Furthermore, with the HHD instruction, the participants made

about twice as many failures as with the HMD instruction. However, statistical significance could not be found between the groups. Therefore, the previously postulated hypothesis - *Do HHD and HMD AR devices differ in their impact on assembly performance (assembly time and failures)?* - can be rejected.

However, by looking into the assembly steps in more detail, the descriptive results reveal that the HMD AR device participants needed less time and made fewer failures. In six of the eight steps, the HMD group was slightly faster than the HHD group. One reason might be that participants using the HMD did not have to put the HMD aside during assembly, as was the case with the HHD device. Nonetheless, this might also be accompanied by the limitation in the perception of the real environment due to the VST technique of the HMD AR device. Compared to competing products, the used HMD AR device is characterized by its good camera quality, which is crucial for the VST AR application. However, compared to the natural perception of the environment, as the test subjects of the HHD group had, participants in the HMD group had perceptible limitations. The camera quality of the glasses was on a level that allowed all participants to assemble the cupboard without major problems. However, the participants in the HMD group showed difficulties with the HMD AR glasses, especially with filigree activities. For example, picking out the right screws, screwing them in, or plugging the side panels together. Nevertheless, on a descriptive level, the HMD group needed less time for the entire assembly process. With further advancements in VST camera quality, HMDs may be able to extend the modest advantages observed in this study, which have thus far been demonstrated only at a descriptive level.

Statistical significance of the group differences regarding failures was achieved in steps three and four. In step 3, the HMD AR application worsened the failure rate compared with HHD ($M_{HHD} = 0$, $M_{HMD} = 0.41$). In step 4, however, the HMD group committed significantly fewer errors than the HHD ($M_{HHD} = 0.75$, $M_{HMD} = 0.09$). In this step, the attached side panels must be screwed to the base plate from the outside (Fig. 2 (d)). The blue screwdriver to be used for this is shown to the right and left of the 3D model of the cupboard. During the experiments, it was frequently observed that participants in the HHD group used the red Phillips screwdriver, which was also used in the previous assembly steps, instead of the blue-slotted screwdriver. One possible explanation is that the participants often overlooked the displayed screwdriver because it was positioned outside the area in which the tablet of the HHD AR application visualizes the virtual objects. The opposite results might be due to a positive effect of the larger FOV by the HMD AR device, which is one advantage the HMD has compared to the HHD.

The advantage of the significantly larger FOV of the HMD AR device could have been measured here.

Another interesting aspect is revealed by a much higher standard deviation in the HMD group, which indicates differences in participants' usability or familiarity with the device. Following experimental studies should consider the prior experience and familiarity participants have with the device in more detail.

4.2 Impact of AR devices on cognitive load

The cognitive load rating, assessed after the assembly process, was nearly similar between the groups, showing an average of 2.44 in the HHD and 2.71 in the HMD group. Statistical significance was not reached. Thus, the investigated research question - *Do HHD and HMD AR devices evoke differences in the user's cognitive load depending on the AR technology used?* - is denied.

In general, the measured results indicate the relatively low complexity of the investigated scenario, which generated a low level of cognitive load ($M_{HHD} = 2.44$, $M_{HMD} = 2.71$ on a 9-point Likert scale) in both groups. The low complexity might be a factor as to why a difference between the AR devices in terms of cognitive load could not be measured. One could assume that the cognitive demands were not at a level that stretched the capacity of the participants. However, the low cognitive load rating also confirms that both AR devices are suitable for providing instructions for a guided assembly process without putting additional strain on the person performing the task. Similarly, the correlation analyses show no significant relationship between cognitive load and performance in the assembly process, which indicates that failures and time were not negatively affected by the cognitive load. Consequently, in the investigated assembly process, making mistakes or slowly assembling the cupboard does not seem to result from a high cognitive load. However, in future studies, we intend to increase the complexity of the assembly scenario to specify our statements.

4.3 Influence of AR devices on motivation

The participants' motivation during the assembly process was measured by four scales of the IMI measurement device covering different aspects of motivation. The results in the IMI scale show no differences in the comparison between the two AR groups, which is why the previous research question - *Do HHD and HMD AR devices affect the motivation of the user?* - is denied.

Results on a subscale level show a high level of satisfaction (subscales *Interest/Enjoyment* and *Perceived Competence*) for both groups during the assembly using the AR devices, which confirms the basic suitability of AR as a transfer instrument in guided assembly tasks. However, it should be noted that the effects on interest and enjoyment may be due to the novelty of the technology and the experimental setting. In a real application scenario in which a person is familiar with the task or the AR device, these effects could diminish. A link between motivational factors and performance in guided assembly tasks could not be measured in this study. The correlation analysis with the *Interest/Enjoyment* scale and each assembly time and failures as dependent variables showed no association. The same was shown for the closely related *Effort/Importance* scale, for which no influence on the success factors could be detected. For the investigated assembly scenario, it can be stated that the motivation of the test persons was also not decisive or detrimental to the number of failures committed or the time required for the assembly. The reason for this could again be the already described low cognitive demand in the assembly tasks. It can be assumed that the assembly to be performed could also be successfully completed with a low level of motivation of the participants due to its low cognitive demands.

4.4 Limitations and outlook

This study investigates the influence of two AR devices on the performance, cognitive load, and motivation in a guided assembly process. First, this study integrates technological, human, and organizational factors since it tests performance (failure and time) and considers cognitive load and motivation important for sustainable usage. Secondly, this study incorporates an experimental setting, which justifies statements about the effectiveness of the two AR devices under controlled conditions using a representative sample.

However, this experimental design was not implemented without limitations. Therefore, the small sample size ($N=33$) and a more precise sample coverage must be considered to differentiate the results. Unlike Grum and Gronau [43], this study did not measure the experience or prior knowledge of assembling an IKEA cupboard (subdivided into amateur and expert). Different results depending on previous experience can be assumed [43]. In addition, measuring other sample-describing factors, such as spatial abilities or experience with AR/VR, or considering potential gender or age differences in the analysis also seems helpful.

Furthermore, considering the cognitive load, results indicate a rather low complexity of the assembly task, resulting in low cognitive demand on the participant's side. A more complex use case, which induces a higher cognitive load of the participants, would be interesting to investigate. In addition, the subjective self-assessed measurement of cognitive load using Paas' scale does not have the potential to capture the complex construct of cognitive load fully. The

assessment could be enhanced through more objective measures such as eye-tracking data or electroencephalography (EEG).

With regard to the comparison of handheld and head-mounted AR displays, it should be noted that only two specific AR devices (Apple iPad and Varjo XR-3) were examined, which entail device-specific technical restrictions. Future research could investigate further AR devices to achieve general conclusions about AR devices.

In addition, other aspects of the user experience for comparison could be considered in future research (e.g., technology acceptance or familiarity). Moreover, the technical capabilities of the devices were not fully exploited. Device-specific advantages of the HMD, such as the possibility of using gaze data or hand recognition as input modalities, could offer a further advantage and are essential for a comprehensive examination. The same applies to the design of the AR application. Different AR devices may need different user interface designs to improve the user experience. Once the device-specific differences of AR devices are identified, there is still a need to design appropriate interfaces to maximize the user experience.

4.5 Summary and implications

The present study investigates the impact of two different AR displays on performance, cognitive load, and motivation in a guided assembly task. Our study findings reveal that Augmented Reality (AR) head-mounted displays (HMD) and AR handheld displays (HHD) are suitable for providing instructions and support in guided assembly tasks, as measured by performance.

This study tested the technology-specific differences between two AR devices in a guided assembly scenario by comparing a VST HHD AR device and a VST HMD AR device in an experimental setting. However, the differences between AR devices and the effects of these differences on their use in various application scenarios have yet to be comprehensively investigated. Due to the diversity of the devices, we encourage researchers to continue conducting experimental research to provide evidence-based recommendations about the practical application of AR in the long run.

To summarize the study's results, both AR devices show nearly similar results in performance, cognitive load and motivation for the given experimental setting. Overall, in both groups, participants made only a few mistakes (up to two mistakes) and reported low cognitive load, which indicates that both AR devices are overall suited for giving instructions in a task with low complexity. However, the technical differences between the VST HMD AR and HHD AR devices are apparent when comparing the AR technologies,

which was a key objective of the study. First, the VST HMD AR device outperforms the HHD AR device's large FOV. However, other device-specific limitations of the Varjo XR-3 limit the full exploitation of the potential. Beyond the limited perception of the environment, the device used is relatively large and heavy compared to other AR devices, which was described as exhausting for some participants throughout the experiment. Especially in assembly tasks, where filigree components must be interacted with, unrestricted perception of reality seems essential for safe execution. Further research could include other HMD AR devices in the investigation (other HMD AR technologies also use VST or OST). It can also be assumed that future technologies will overcome the current technical limitations, requiring constant re-evaluation of HMD AR. This knowledge about the impact of the technologies in the application is crucial to exploit the full potential of AR.

Declarations

References

1. Panzer M, Bender B (2022) Deep reinforcement learning in production systems: a systematic literature review. Int J Prod Res 60:4316–4341. https://doi.org/10.1080/00207543.2021.1973138
2. Sautter B, Daling L (2021) Mixed reality supported learning for Industrial on-the-job training. SSRN J. https://doi.org/10.2139/ssrn.3864189
3. Teichmann M, Lettkemann V, Gronau N (2024) Digitalization, demographic change and decarbonization: eight pivotal competencies for learning factories. In: Thiede S, Lutters E (eds)

Learning factories of the future. Springer Nature Switzerland, Cham, pp 313–320

4. Gladysz B, Tran T, Romero D, van Erp T, Abonyi J, Ruppert T (2023) Current development on the Operator 4.0 and transition towards the Operator 5.0: a systematic literature review in light of industry 5.0. J Manuf Syst 70:160–185. https://doi.org/10.1016/j.jmsy.2023.07.008

5. Howard MC, Davis MM (2023) A Meta-analysis of augmented reality programs for education and training. Virtual Reality. https://doi.org/10.1007/s10055-023-00844-6

6. Daling LM, Schlittmeier SJ (2024) Effects of Augmented Reality-, virtual Reality-, and mixed reality–based training on Objective Performance measures and subjective evaluations in Manual Assembly tasks: a scoping review. Hum Factors 66:589–626. https://doi.org/10.1177/00187208221105135

7. Azuma RT (1997) A survey of augmented reality. Presence: Teleoperators Virtual Environ 6(4):355–385. https://doi.org/10.1162/pres.1997.6.4.355

8. Wu H-K, Lee SW-Y, Chang H-Y, Liang J-C (2013) Current status, opportunities and challenges of augmented reality in education. Comput Educ 62:41–49. https://doi.org/10.1016/j.compedu.2012.10.024

9. Yu Z (2023) Retracted article: Meta-analyses of effects of augmented reality on educational outcomes over a decade. Interact Learn Environ 0:1–15. https://doi.org/10.1080/10494820.2023.2205899

10. Lin Y, Yu Z (2023) A meta-analysis of the effects of augmented reality technologies in interactive learning environments (2012–2022). Comput Appl Eng Educ 31:1111–1131. https://doi.org/10.1002/cae.22628

11. Buchner J, Buntins K, Kerres M (2021) The impact of augmented reality on cognitive load and performance: a systematic review. J Comput Assist Learn jcal 12617. https://doi.org/10.1111/jcal.12617

12. Ozdemir M, Sahin C, Arcagok S, Demir MK (2018) The effect of augmented reality applications in the learning process: a Meta Analysis Study. EJER 18:1–22. https://doi.org/10.14689/ejer.2018.74.9

13. Kappler E, Jacinto RF, Arndt S (2022) Evaluation of Visual Acuity and Perceptual Field of View Using the Varjo XR-3 Headset in a Virtual Environment. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 66:2193–2197. https://doi.org/10.1177/1071181322661195

14. Bläsing D, Bornewasser M, Hinrichsen S (2021) Informatorische Assistenzsysteme. In: Bauer W, Mütze-Niewöhner S, Stowasser S, Zanker C, Müller N (eds) Arbeit in Der Digitalisierten Welt: Praxisbeispiele Und Gestaltungslösungen aus dem BMBF-Förderschwerpunkt. Springer, Berlin, Heidelberg, pp 257–271

15. Baird KM, Barfield W (1999) Evaluating the effectiveness of augmented reality displays for a manual assembly task. Virtual Reality 4:250–259. https://doi.org/10.1007/BF01421808

16. Tang A, Owen C, Biocca F, Mou W (2003) Comparative effectiveness of augmented reality in object assembly. New Horiz 8. https://doi.org/10.1145/642611.642626

17. Hou L, Wang X, Bernold L, Love PED (2013) Using animated augmented reality to cognitively Guide Assembly. J Comput Civ Eng 27:439–451. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000184

18. Yang Z, Shi J, Jiang W, Sui Y, Wu Y, Ma S, Kang C, Li H (2019) Influences of augmented reality assistance on performance and cognitive loads in different stages of Assembly Task. Front Psychol 10:1703. https://doi.org/10.3389/fpsyg.2019.01703

19. Havard V, Baudry D, Jeanne B, Louis A, Savatier X (2021) A use case study comparing augmented reality (AR) and electronic document-based maintenance instructions considering tasks complexity and operator competency level. Virtual Reality 25:999–1014. https://doi.org/10.1007/s10055-020-00493-z

20. Kolla SSVK, Sanchez A, Plapper P (2021) Comparing effectiveness of paper based and augmented reality instructions for manual assembly and training tasks. In: Proceedings of the Conference on Learning Factories (CLF)

21. Mourtzis D, Zogopoulos V, Xanthi F (2019) Augmented reality application to support the assembly of highly customized products and to adapt to production re-scheduling. Int J Adv Manuf Technol 105:3899–3910. https://doi.org/10.1007/s00170-019-03941-6

22. Funk M, Kosch T, Schmidt A (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. ACM, Heidelberg Germany, pp 934–939

23. Drouot M, Bigot NL, Bricard E, de Bougrenet J-L, Nourrit V (2022) Augmented reality on industrial assembly line: impact on effectiveness and mental workload. Appl Ergon 103:103793. https://doi.org/10.1016/j.apergo.2022.103793

24. Schein KE, Rauschnabel PA (2023) Augmented reality in Manufacturing: exploring workers' perceptions of barriers. IEEE Trans Eng Manage 70:3344–3357. https://doi.org/10.1109/TEM.2021.3093833

25. Paas FGWC, Van Merriënboer JJG (1994) Instructional control of cognitive load in the training of complex cognitive tasks. Educational Psychol Rev 6:351–371. https://doi.org/10.1007/BF02213420

26. Sweller J, Van Merrienboer JJ, Paas FGWC (1998) Cognitive architecture and instructional design. Educational Psychol Rev 10:251–296

27. Paas FGWC, van Merriënboer JJG, Adam JJ (1994) Measurement of cognitive load in Instructional Research. Percept Mot Skills 79:419–430. https://doi.org/10.2466/pms.1994.79.1.419

28. Suzuki Y, Wild F, Scanlon E (2024) Measuring cognitive load in augmented reality with physiological methods: a systematic review. J Comput Assist Learn 40:375–393. https://doi.org/10.1111/jcal.12882

29. Kosch T, Karolus J, Zagermann J, Reiterer H, Schmidt A, Woźniak PW (2023) A survey on measuring cognitive workload in Human-Computer Interaction. ACM Comput Surv 55:283:1–283. https://doi.org/10.1145/3582272

30. Carmigniani J, Furht B, Anisetti M, Ceravolo P, Damiani E, Ivkovic M (2011) Augmented reality technologies, systems and applications. Multimedia Tools Appl 51:341–377. https://doi.org/10.1007/s11042-010-0660-6

31. Bottani E, Vignali G (2019) Augmented reality technology in the manufacturing industry: a review of the last decade. IISE Trans 51:284–310. https://doi.org/10.1080/24725854.2018.1493244

32. Alves J, Marques B, Oliveira M, Araujo T, Dias P, Santos BS (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). IEEE, Porto, Portugal, pp 1–6

33. Alves JB, Marques B, Ferreira C, Dias P, Santos BS (2022) Comparing augmented reality visualization methods for assembly procedures. Virtual Reality 26:235–248. https://doi.org/10.1007/s10055-021-00557-8

34. Blattgerste J, Strenge B, Renner P, Pfeiffer T, Essig K (2017) Comparing Conventional and Augmented Reality Instructions for Manual Assembly Tasks. In: Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments. ACM, Island of Rhodes Greece, pp 75–82

35. Büttner S, Funk M, Sand O, Röcker C (2016) Using Head-Mounted Displays and In-Situ Projection for Assistive Systems: A Comparison. In: Proceedings of the 9th ACM International

Conference on PErvasive Technologies Related to Assistive Environments. ACM, Corfu Island Greece, pp 1–8

36. Ragan ED, Bowman DA, Kopper R, Stinson C, Scerbo S, McMahan RP (2015) Effects of Field of View and Visual Complexity on virtual reality training effectiveness for a visual scanning Task. IEEE Trans Visual Comput Graphics 21:794–807. https://doi.org/10.1109/TVCG.2015.2403312

37. Baumeister J, Ssin SY, ElSayed NAM, Dorrian J, Webb DP, Walsh JA, Simon TM, Irlitti A, Smith RT, Kohler M, Thomas BH (2017) Cognitive cost of using augmented reality displays. IEEE Trans Visual Comput Graphics 23:2378–2388. https://doi.org/10.1109/TVCG.2017.2735098

38. Ren D, Goldschwendt T, Chang Y, Hollerer T Evaluating wide-field-of-view augmented reality with mixed reality simulation. In: 2016, Reality IEEEV (2016) (VR). IEEE, Greenville, SC, USA, pp 93–102

39. Trepkowski C, Eibich D, Maiero J, Marquardt A, Kruijff E, Feiner S (2019) The Effect of Narrow Field of View and Information Density on Visual Search Performance in Augmented Reality. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). IEEE, Osaka, Japan, pp 575–584

40. Niegemann HM, Heidig S (2012) Multimedia Learning. In: Seel NM (ed) Encyclopedia of the sciences of Learning. Springer US, Boston, MA, pp 2372–2375

41. Moreno R, Mayer R (2007) Interactive Multimodal Learning environments: special issue on interactive learning environments: Contemporary issues and trends. Educ Psychol Rev 19:309–326. https://doi.org/10.1007/s10648-007-9047-2

42. Blumenfeld PC, Kempler TM, Krajcik JS (2006) Motivation and Cognitive Engagement in Learning environments. The Cambridge handbook of: the learning sciences. Cambridge University Press, New York, NY, US, pp 475–488

43. Grum M, Gronau N (2020) Adaptable knowledge-driven Information systems improving knowledge transfers. In: Shishkov B (ed) Business modeling and Software Design. Springer, Cham, pp 202–220

44. Grum M (2022) Welcome to the Platform for Augmented Reality-based Learning and Knowledge Transfers. https://github.com/MarcusGrum/augmented-reality-based-learning. Accessed 6 Sep 2024

45. Paas FGWC (1992) Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach. J Educ Psychol 84:429–434. https://doi.org/10.1037/0022-0663.84.4.429

46. Center for Self-Determination Theory (2023) Intrinsic Motivation Inventory (IMI). https://selfdeterminationtheory.org/intrinsic-motivation-inventory/. Accessed 6 Sep 2024

47. Kendall MG (1955) Rank correlation methods. Griffin, London