Management and Production Engineering Review
Volume 15 • Number 1 • March 2024 • pp. 12–34
DOI: 10.24425/mper.2024.149987

The Impact of Augmented Reality Devices on Operator Performance in Manufacturing Contexts

Chiara BELSITO[1], Chiara FRANCIOSI[2], Valentina DI PASQUALE[1]

[1] *Department of Industrial Engineering, University of Salerno, Italy*
[2] *Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France*

Received: 14 April 2023
Accepted: 16 November 2023

Abstract
This study explores the impact of augmented reality (AR) on worker performance in manufacturing contexts through an analysis of case studies extant in the literature. Two specific analyses were conducted to assess the impacts of AR technologies on worker performance in terms of objective and subjective metrics, and in terms of their age, experience with the task and experience with the AR device. Regarding objective metrics, the results showed that the task completion time was reduced for some AR devices (projectors, monitors, tablets, smartphones), whereas the use of the head-mounted display (HMD) increased task-completion time; moreover, the error rate was reduced with any AR device compared with traditional methods. Regarding subjective metrics, the analysis underlined that operator perceived a lower workload with the HMD or the monitor compared with traditional methods. The age of operators did not influence performance, while the operators' experience allowed for the improvement of human performance.

Keywords
Human factors, human errors, industry 4.0, manufacturing systems; technology implementation, smart factory.

Introduction

The industry 4.0 (I4.0) paradigm encourages the digitization and connection of equipment in manufacturing plants, revolutionising workplaces: big data, the Internet of Things (IoT), high-speed connectivity, artificial intelligence (AI), human–machine interaction and the digital twins of physical assets are the essential elements destined to radically change the manufacturing sector. In this complex context, humans will play a central role, and new types of interactions between humans and machines will generate significant implications for the nature of the work (Romero et al., 2016; Longo et al., 2017; Romero et al., 2020). The introduction of new technologies enabling Industry 4.0 can help operators become smarter in some activities, but it can also have a negative impact on human and business performance; the interaction between different technologies and humans influences quality, safety, and productivity. The extent of this influence depends on the variability of several operator characteristics, such as age, experience, reliability, and behaviour. Among several 4.0 technologies, augmented reality (AR) is widely adopted by operators in many activities, such as the design phase as well as the production, assembly and maintenance processes, in order to enhance industrial applications from an ergonomic and economic point of view (Di Pasquale et al., 2022; Sharma et al., 2022). Indeed, AR technology provides information and digital data superimposed in real time in the user's field of sight (e.g., tablets, smartphones, AR space projectors), enriching the real factory environment of operators.

Although some literature reviews have investigated the relationship between AR technology and operators in different contexts, such as logistics, production, maintenance, and several case studies have aimed at evaluating AR implementation in traditional working operations (Palmarini et al., 2018; Bottani and Vignali, 2019; Danielsson et al., 2020; Lagorio et al., 2022), an overall assessment of the AR impacts on human performance has not been conducted. A previous review (Di Pasquale et al. 2022) has underlined the need to focus on AR technologies by thoroughly evaluating the impacts of varying devices used, to identify and

***Corresponding author**: Valentina Di Pasquale – University of Salerno, Department of Industrial Engineering, 84084, Fisciano, Italy, e-mail: vdipasquale@unisa.it*

define the most suitable fields of application and to analyse the effect on the performance of operators of factors such as experience, age, trust, and acceptance.

For this reason, the goal of our study was to analyse the impact of the use of AR on the performance of operators in manufacturing contexts through the screening and comparison of all pertinent case studies found in the scientific literature. In particular, the analysis was conducted to answer the following research questions (RQs):

- RQ1. How do the different types of AR devices impact worker performance?
- RQ2. How do specific AR devices impact the worker performance, taking human 'age' and 'experience' into account?

Age and experience were chosen among the operator attributes to be analysed in depth in this study as they are objectively measurable and identifiable, as opposed to subjective attributes such as trust and acceptance of technology. Therefore, first, the information related to the operators (i.e., the attributes characterising the human performance) and the objective and subjective metrics affected using AR were extracted from the case studies. Next, the positive and negative impacts of AR on these metrics (task completion time, error rate, perceived workload, usability, learning rate and others) were analysed considering various factors characterising human performance (i.e., age, experience, and the type of AR device investigated). This paper is structured as follows. Section 2 provides the method adopted for reaching the goal of the study. Section 3 reports the results of the conducted analysis, and Section 4 presents an in-depth discussion of the results, the conclusions, and the identified research gaps.

Materials & Methods

First, a systematic literature review was conducted. The classification criteria for the content analysis of the papers selected through the literature review are identified and described in following sections, including the description of the type of analysis conducted to answer the research questions.

Systematic literature review

A systematic literature review was conducted consecutively in the following four steps:

1. Definition of suitable keywords.
2. Search for papers in the selected database under appropriate limitations.
3. Selection of papers according to screening criteria.
4. Analysis of selected papers and data extraction.

We used the Scopus database to conduct the search. The keywords were divided into four groups (Table 1): the first, 'technology', which concerns the various synonyms with which AR can be expressed in order to cover as many papers as possible published in the literature on the topic; the second, 'sector', which concerns the various terms with which the production contexts can be expressed in order to limit research to the industrial sector; and the third, 'operator', which refers to the workforce and 'performance', which refers to the performance of the operator that might be affected by the introduction of AR into the manufacturing system. All possible combinations of keywords from the groups were considered using the Boolean 'AND' operator between each group and the 'OR' operator within each group. The asterisk was used as a truncation function that allowed searching for all variants of a word. Searches covered the titles, keywords, and abstracts of papers in the database.

Table 1
List of keywords selected for the systematic literature review

	Keywords			
	Technology	Sector	Operator	Performance
← OR →	AR; Augmented reality; Computer-mediated reality.	Manufact*; Industr*; Factory; Production.	Worker; Operator; Employee; Ageing workforce.	Performance; Productivity; Efficiency; Safety; Error; Quality; Time; Reliability; Effort; Workload.
	← AND →			

Regarding the restriction criteria, the review was limited to articles published in the last 10 years (2011–2021) and written in the English language in the following subject areas: computer science, engineering, mathematics, decision sciences, social sciences, business management and accounting, material science, economics, econometrics and finance, neuroscience, and psychology. Selection screening was divided into two steps. The first screening involved reading the titles and abstracts of each paper and excluding papers according to the following criteria:

- Papers in which AR was analysed only from a technological point of view and not from the operator's point of view. State-of-the-art analyses, systematic literature reviews and research and surveys on the

various enabling technologies introduced by Industry 4.0 were also considered in this first exclusion criterion.
- Papers not related to the production sector, but which referred to healthcare, construction, maritime, aeronautical, aerospace, energy and agriculture applications.
- Papers related to virtual reality or mixed reality.

The second screening involved reading the full text of the papers selected in the first step which represented the final assessment of the most relevant documents. In this step, the paper sample was analysed further, and other articles were excluded for the following reasons: the full text was not available; the features of the sample selected for the case study were missing or not sufficiently detailed; the type of device used was not specified; and papers describing case studies carried out in the company or in a laboratory, but not reporting the obtained numerical results.

Classification criteria for the content analysis

After a full-text reading of the selected papers, the classification criteria were defined (Table 2), and in accordance with them, the relevant information was extracted and stored on an Excel spreadsheet to facilitate data analysis.

The selected case studies, each with a specific ID, were further analysed, extrapolating information about the operator and objective and subjective metrics. To assess the impacts of AR on operator performance, a distinction was made between objective and subjective metrics.

Objective metrics included the following:
- Task completion time, that is, the time taken by operators to complete a given task.
- Error rate, which identifies the number of mistakes made.

Subjective metrics included the following:
- Perceived workload evaluated through the NASA Task Load Index (NASA-TLX) indicator or RAW Task Load Index (RAW-TLX) version.
- Usability evaluated through the system usability scale (SUS) indicator.
- Learning rate, which allows assessment of whether the operator with the AR can learn the instructions with respect to traditional methods.
- Qualitative assessments based on interviews or surveys that were administered to operators.

For each case study, whether the individual's performance was impacted positively (↗: improved performance), negatively (↘: worsened performance) or in a non-significant way (→: i.e., carrying out the activity with or without AR was indifferent) was analyzed.

Table 2
Classification criteria for the content analysis

Criteria	Definition
Bibliometric indicators	Author; Type of document (Journal or Conference); Year of publication; Source title; Keywords.
Type of contribution	Papers were classified including different types of contribution: development of a solution or case study for evaluating impacts of technology.
Sample	Description of the sample of selected operators to carry out case study, reporting information regarding their age and experience.
Research environment	The research environments of the case study were laboratory, industry, or both.
AR devices	AR devices were distinguished among Head Mounted-Display (HMD), projector, tablet, monitor and smartphone with their features.
Type of comparison	Type of comparison on which the evaluation of impacts on human performances was based. The selected papers were divided in (1) studies that compared the worker performance with and without the implementation of one of the AR devices and (2) studies that compared the impacts of different AR devices on operators.
Type of application	Types of applications were assistance, training, remote collaboration and also the combination of two types of application.
Type of task	Types of tasks were assembly, maintenance, quality control, health & safety, design variations, order picking, packaging, pick and place.
Subjective metrics	Types of qualitative metrics are perceived workload, usability, learning rate and other qualitative evaluations.
Objective metrics	Types of quantitative metrics are task completion time and error rates.

These conclusions were defined according to the results of the statistical analyzes (ANOVA, MANOVA, Fisher's test, t test, variance analysis, etc.) that each case study reported for both objective and subjective metrics. Further, it was possible to identify the total number of papers that evaluated a certain metric, regardless of the type of comparison between the different AR devices analyzed in the case study.

Analysis of the AR impacts on the worker performance

To satisfy the goal of this study and to answer the two RQs, two specific analyses were conducted.

The first analysis aimed to assess the impacts of AR on worker performance with respect to the type of AR device used. The device impacts were compared in pairs, considering all types of AR devices (HMD, projector, monitor, smartphone, and tablet) and traditional methods, including paper or digital supports used by the operators to carry out a specific task. For each pairwise comparison, there were three options for a given metric: device X achieves better or worse or the same result as device Y. For example, by comparing the HMDs and projectors with respect to task completion time (objective metric), this showed which of the two devices allowed the operator to complete the task faster, or by comparing the HMDs and the traditional method with respect to the error rate (objective metric), this indicated which of the two solutions allowed the operator to commit fewer errors.

The second analysis aimed to assess the impacts of the specific AR device on worker performance by varying their attributes. Regarding the operators, the attributes identified are shown in Table 3.

Table 3
Human attributes

Attributes	Definition
Age	Age can be expressed as average age or an age range.
Experience with AR	Experience with AR to determine if the operators had previous experience with the devices or if they were novices.
Experience with the task	Experience with the task to determine whether the operators had previous experience with this activity, or if it was their first exposure.

In this analysis, the study focused only on papers for which information was available regarding level of operator experience with AR to draw conclusions related to the links to AR devices, age and experience. Only the comparisons between the AR device and the traditional methods were considered. The case studies were analysed by level of experience – high/medium/low or no experience- based on the classification made by the authors of the selected papers. For each level of experience, the results were analysed to compare a single type of AR device. In analysing the case studies, the age of the involved participants was also considered. This was done to determine how the impacts on the performance of operators varied, depending not only on the type of AR device used, but also taking the level of experience and age into account. The analysis was conducted for all metrics, both objective and subjective, to assess the impacts, as the level of experience and age of the users varied.

Results

SLR results

Figure 1 shows the results of the systematic literature review process and the number of papers selected for each step. Combining the defined keywords, the search database found 754 papers.

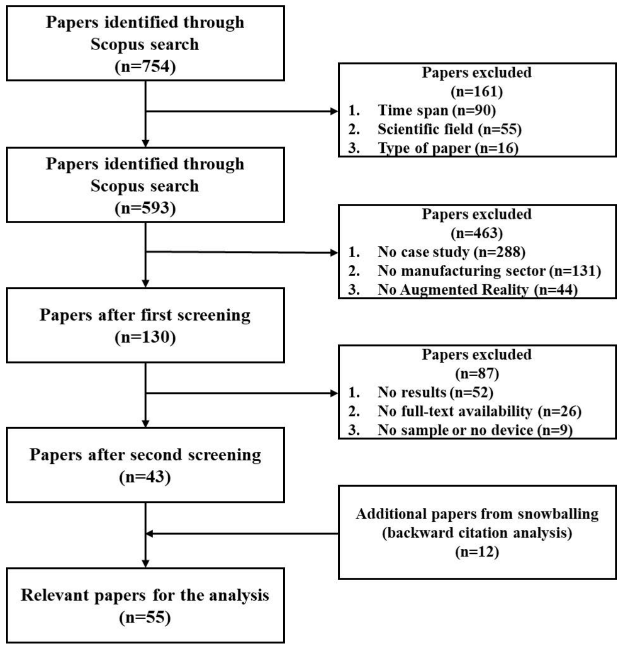

Fig. 1. SLR results

Selecting the last 10 years (2011–2021) as the timeline; the scientific field as the topic; the scientific journal, conference proceeding, book chapter or review as the type of document; and English as the language, the total number of studies was reduced to 593. The 593 selected papers were analysed, and several were excluded according to previously defined criteria. Following the selection screening process, after the title and abstract reading (first step), 288 papers were excluded because they did not report any case study, 131 because they referred to a different sector than production, and 44 because they dealt with virtual or mixed reality. Therefore, in the first screening, only

130 papers were deemed relevant. After the full-text reading of papers (second step), the sample of eligible papers was reduced to 43 papers, excluding 26 articles because the full text was not available, 9 because the sample of operators selected or the type of device used was not specified, and 52 because they were incomplete studies. The other 12 papers from the snowball process, i.e., the reference screening of the selected papers, were finally added, resulting in a sample of 55 studies (Appendix A).

As underlined in Figure 2, the number of selected documents grew over time, and most of the identified papers had been published in recent years. Among the selected papers, 35 were published in conference proceedings and 20 in scientific journals.

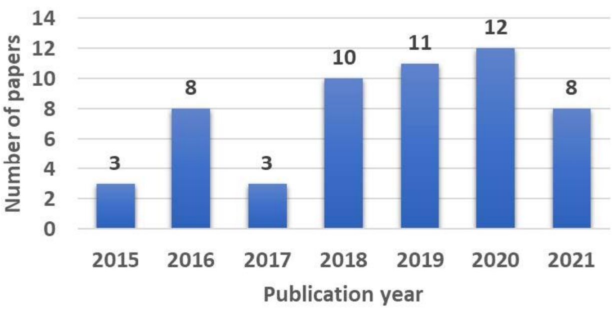

Fig. 2. Publications per year

Results of the content analysis

Considering the 'research environment' criterion, 33 case studies were carried out in a laboratory; the remaining 22 were carried out directly in an industrial context.

The identified AR devices were classified into four groups and listed in descending order, from most used to least used, according to the obtained results:

1. HMDs were the most frequently used within the examined case studies. Their hands-free feature allows the operator to maintain eye contact with the activity that they are conducting, significantly limiting distractions resulting from the use of the device. It is important for a successful hands-free system that virtual content is perceived by the user in the most natural way possible so that the mix of real and augmented information is processed by the human brain as a single coherent cognitive flow.
2. Projectors are considered a valid alternative to HMDs as they allow the information to be projected directly in the workplace without wearing any viewer on the head.
3. The third group includes handheld devices, i.e., smartphones and tablets. Tablets are easy to implement and inexpensive, but in practice, operators must employ one or both hands to view information, limiting their ability to operate them. The smartphone is never the first choice in industrial applications because their small screen does not allow a wide enough view of the required data, and in some situations, the size and fragility of the device make it difficult to use.
4. Monitors represent the least used solution because they distract the operators, who have to look away from the activity to capture the information shown on a display.

Of the types of AR devices, HMDs (commonly called smart glasses) were the most adopted, followed by projectors, smartphones, tablets, and a lesser number considered monitors (Fig. 3). Figure 4 reports the trend of the different AR devices over the years. The main device used from 2015 to 2017 was the projector; however, from 2017 to 2021, HMD use clearly prevailed. Comparing the trend of the smartphone with that of the tablet, before 2017, the tablet was the most used of the two even before 2016, but no case study has yet investigated the use of the smartphone with AR in manufacturing contexts. During 2016, the smartphone curve gradually began to rise, and from 2017 to 2021, the curves were inverted. Over this time, the use of smartphones prevailed over the tablet, while the trend of tablets remained constant. The monitor was always the least used solution starting in 2017, except for 2020. As of 2021, the smartphone curve showed an increasing trend, while that of the HMDs, projectors and monitors exhibited a decreasing trend.

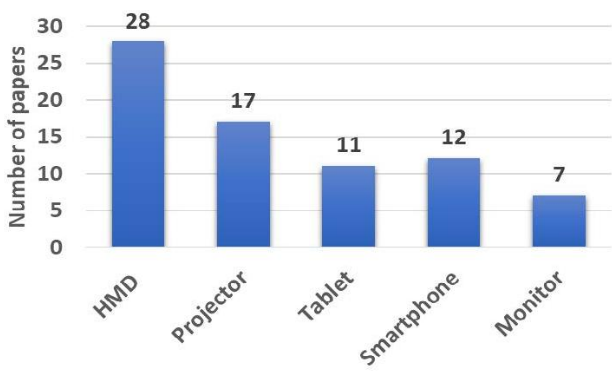

Fig. 3. Type of AR device

In respect to the specific type of task, Figure 5 shows the clear prevalence of AR use for assembly, followed by maintenance, quality control, order picking, health and safety, design variations, packaging and pick and place. Furthermore, the paper analysis highlighted the

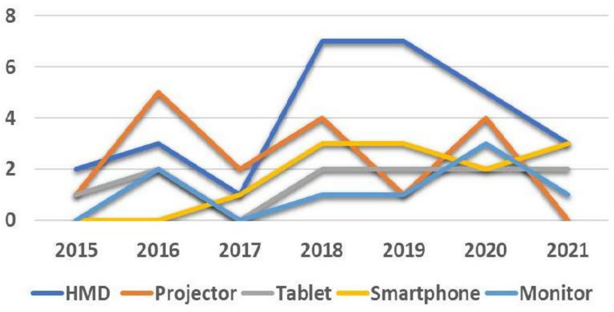

Fig. 4. Trend of devices over the years

prevalence of AR technology to support operators and guide them in their paths (assistance task 21 papers, 38% of the sample), then AR technology to provide instructions to them to accelerate their learning phase (14 papers, 22% of the sample), 12 papers jointly assistance and training and finally AR is used to encourage remote collaboration between an expert operator and an inexperienced one, without the former necessarily being present in the workplace (8 papers). Based on the obtained results, a further cross-analysis was carried out to better understand how the various criteria related to each other. Figure 6 shows which tasks were investigated over the years. The tasks were distributed equally from 2015 to 2021. In particular, in 2015 there were assembly, order picking and quality control, while in 2021 there were assembly, health and safety, maintenance and quality control. The assembly task showed an increasing trend over the years, which culminated in 2020 with nine case studies. However, in the last three years, the use of AR for maintenance was constant.

Fig. 5. Type of tasks

This technology was also tested for quality control, and given the detected benefits, its use doubled in 2021. Figure 7 reports which types of AR devices were primarily used for each individual task.

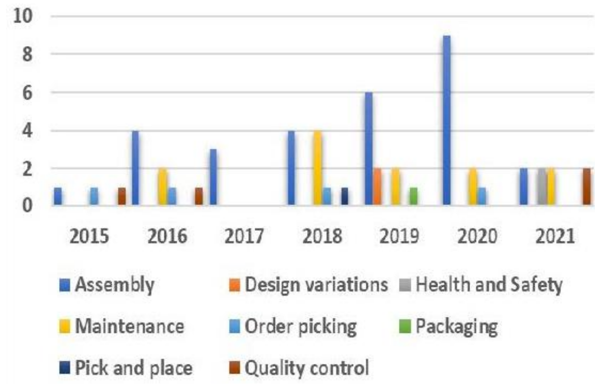

Fig. 6. Trend of tasks over the years

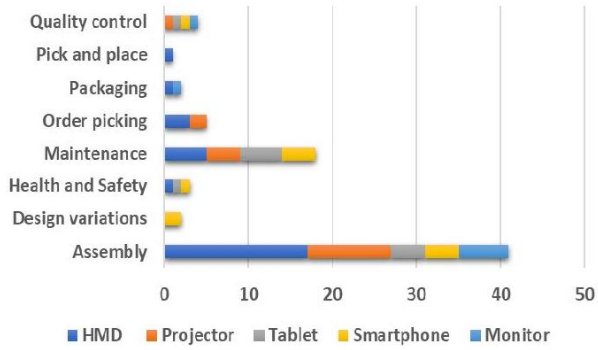

Fig. 7. Relationship between task and device

Concerning the assembly activity, the use of smart glasses clearly prevailed, followed by projectors and finally by tablets, smartphones and monitors. With respect to the activity of design variations (i.e., through an AR device, it is possible to verify whether the arrangement of the components is coherent with the design), there was exclusive use of the smartphone, while for health and safety, there was fairly equal use among HMDs, tablets and smartphones. Smart glasses and tablets, followed by projectors and smartphones, were the devices mostly used for maintenance activities, whereas HMDs, followed by projectors, were the devices generally adopted for order picking. For packaging activity, there was a fairly equal use of smart glasses and monitors, while in the activity of pick and place, HMD was the most widely adopted device. Finally, all types of devices, except HMD, were used for quality control. In Figure 8, the criterion 'type of task' is evaluated based on the 'research environment' to identify which research environment was preferred for each type of task. All types of devices selected in the case studies were tested mainly in a laboratory environment.

Figure 9 shows how the research environment varied over time. From 2015–2020, laboratory experiments prevailed over those in industry.

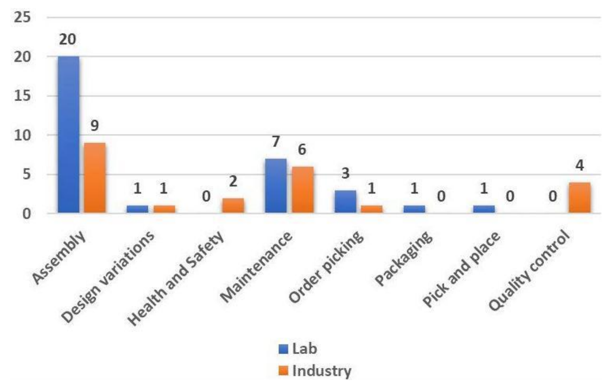

Fig. 8. Relationship between research environment and AR device

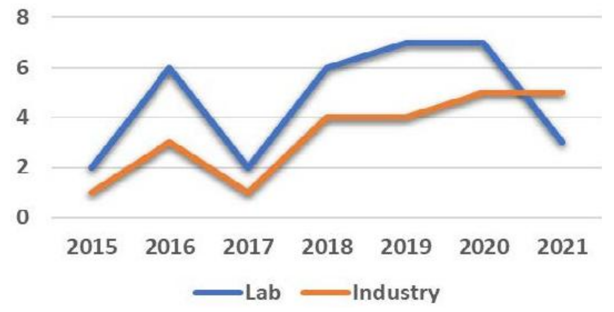

Fig. 9. Trend of the research environment over the years

To identify which type of metric was the most evaluated in the case studies, the objective and subjective metrics were analysed separately, also considering the types of devices that were compared in the various experiments. In Figure 10, the blue colour represents all the types of comparisons in pairs among the devices done in the 55 analysed case studies with respect to the metrics (i.e., one device was compared with another device to assess the impact of their use on one metric). The orange colour represents the comparisons between a specific AR device and the traditional method with respect to the metric. For example, for the objective metric 'task completion time,' a total of 80 comparisons were identified, and a subset of 56 had, as the object of comparison, an AR device and the traditional method. Objective metrics were more common than subjective metrics, having a similar presence for both task completion time and error rate. Among the subjective metrics, perceived workload by the operator prevailed, followed by usability, qualitative evaluation and, finally, the learning rate.

By focusing only on the case studies that evaluated the comparison between any AR device and the traditional method, we found that the most evaluated metric was task completion time, followed by error rate, perceived workload, usability, qualitative evaluation and learning rate.

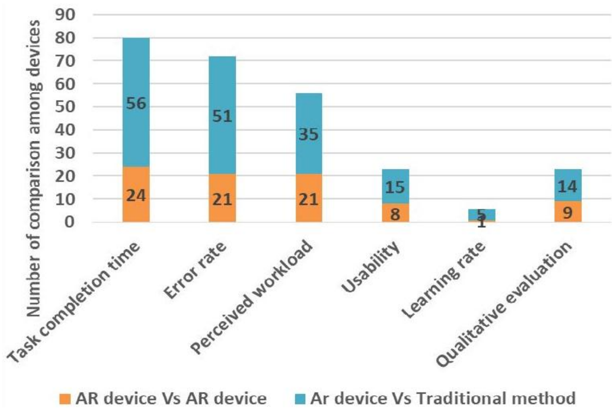

Fig. 10. Analysis of objective metrics vs subjective metrics

AR impacts on the worker performance

Table 4 show the results from the first analysis in which the impacts on worker performance were assessed according to the AR device used. For each metric, the results are presented, taking into consideration only the comparison between AR devices and traditional methods, which included the support provided to operators by both paper and digital manuals. The results of the case studies found in the literature (Appendix B) were used to evaluate the improved/worsened performance.

Concerning task completion time, a negative result was attributed to HMDs, i.e., the operator took more time to complete the task with this device than with traditional methods; this may have been due to the operator's unfamiliarity with the HMDs or they had limited/no experience with the developed AR solution (IDs 2-18-27-30-31-44-48-55). Instead, the projector, tablet, monitor and smartphone were reported as producing positive results; that is, the operator took less time to complete the activity (IDs 1-3-8- 9-11-14-15-20-22-25-27-28-35-40-41-45-53). With respect to the error rate, the HMD, projector, tablet and monitor were more advantageous compared with traditional methods, i.e., the operator made fewer errors using AR technology (IDs 1-2-5-9-11-16-18-20-22-23-26-27-29-30-31-37-39-40-42-43-47-52-53-55). The smartphone had an error rate similar to that of traditional methods (IDs 14-15-46-48) – its use did not significantly improve human performance. Concerning perceived workload, the HMD and the monitor allowed the operator to perceive a lower workload than traditional methods (IDs 6-9-16-25-29-30-32-37-43); a workload almost equal to that of traditional methods was associated with the projector and the tablet (IDs 1-10-22-36-47-51). With respect to usability, the HMD, projector and tablet achieved a higher score than with traditional methods (IDs 5-9-10-12-14-16-18-29-37). Concerning learning

Table 4
Results according to the AR device used (first analysis)

	Metrics						
	Objective		Subjective				
	Task completion time	Error rate	Perceived workload	Usability	Learning rate	Qualitative evaluation	
AR device: HMD	↘	↗	↗	↗	→	↘	Traditional method
AR device: Projector	↗	↗	→	↗	–	↗	Traditional method
AR device: Tablet	↗	↗	→	↗	–	↗	Traditional method
AR device: Monitor	↗	↗	↗	–	–	–	Traditional method
AR device: Smartphone	↗	→	–	–	–	–	Traditional method

*Legend: ↗: improved performance, ↘: worsened performance, →: no significant correlation

rate, the operator was able to learn the instructions at almost the same speed, whether they used the HMD or the traditional method (IDs 16-18); therefore, the use of AR technology did not significantly improve human performance. With respect to qualitative evaluation, HMDs were negatively evaluated by workers because some smart glasses limit the operator's view or are bulky and heavy (IDs 41-46); it is difficult to wear them for a full 8-hour shift.

According to some interviewees, HMDs have a low brightness and are not suitable on surfaces in intense light conditions (e.g., sunlight) (ID 44). However, a positive evaluation was given to the projector and the tablet by the operators. The projector allows the operator to project the instructions directly onto the workspace without having to wear a viewer or to search for information in paper manuals. The tablet allows the operator to immediately view the information if they are already familiar with the device (IDs 9-20-22-24-26-41).

After identifying the different metrics analysed in the papers, the case studies were classified according to the attributes of the operators, i.e., age, experience with the AR and experience with the task for the second analysis. Based on the information about the operators described in the selected papers, four ranges were defined: 1) the age of operators less than 35 years, (2) between 36 and 50, (3) more than 50, and (4) 'not classified' when the age of the operator was not reported. The results are reported in Table 5.

Table 5
Classification of papers according to the attributes

Attributes	Range	Number of papers
Age	$a \leq 35$	24
	$36 \leq a \leq 50$	8
	$a > 50$	0
	Not classified	14
Experience with AR	High	2
	Medium	3
	Low or No Experience	25
Experience with the task	High	6
	Medium	6
	Low or No Experience	18

The age (a) of the operators was classified according to four different ranges, i.e., $\leqslant 35$, $36 = 50$, > 50 and not classified. With respect to experience with the AR, a further classification was made depending on the level of familiarity that the operators had with the AR technology. The results show that, overall, most of the employees were under 35 years of age, and they had limited or no knowledge of this technology and of the performed task. Although a greater number of articles present information on the experience with the task of the operators involved in the experimentation, the notable differences between the type of task performed and the lack of results differentiated by level of experience with the task in the selected papers did not allow for in-depth analysis of this aspect. While instead, albeit with a more limited number, it was possible to evaluate the effects of the experience with AR technology.

Indeed, based on the attributes of the samples, a combined analysis of human performance at varying levels of the experience with AR technology, the age of the operator and the type of the AR device used to carry out the task was conducted. To achieve this goal, the papers were collected and sorted by level of experience (see Table 6). For each, the age of the operators, the type of AR device used and the impacts on both objective and subjective metrics were determined. In this way, it was possible to analyse how the impacts on the operator performance varied, depending not only on the type of AR device used but also on level of experience and age. With respect to a high level of experience with AR, there were two papers available (IDs 37-39) in which only two types of comparison, 'HMD vs Traditional method' and 'Monitor vs Traditional method', were examined. With an average level of experience, three papers were identified (IDs 9-12-14), for which, by analysing the information contained within them, the results were: two comparisons of 'Tablet vs Traditional method', two comparisons of 'HMD vs Traditional method' and 1 comparison of 'Smartphone vs Traditional method'. Regarding a low level of experience, 25 papers were identified (IDs 2-5-13-15-18-20-24-26-28/33-40/44-47-48-50-54-55), of which there were three comparisons of 'Tablet vs Traditional method', eight comparisons of 'HMD vs Traditional method', four comparisons of 'Projector vs Traditional method' and two comparisons of 'Smartphone vs Traditional method'.

For the 'HMD vs Traditional method' comparison, the literature provided at least one case study for each level of experience. The results are summarised in Table 7. When the operator had high experience with AR and was aged 20 to 40, the use of the AR device provided benefits in completion time, error rate and perceived workload rather than usability. When the AR experience reached an average level and the operators were from 19 to 46 years of age, the impacts on performance were the same as those with a high level of

Table 6
Classification of the papers based on level of experience/age and impacts on objective/subjective metrics

Experience with AR	Age	Device	Metrics				IDs
			Task completion time	Error rate	Perceived workload	Usability	
High	20–40; 23–40	HMD vs Traditional method	– ↗	↗ –	↗ –	↗ –	37, 39
High	20-40	Monitor vs Traditional method	–	↗	↗	↗	37
Medium	21–62; 24,2	Tablet vs Traditional method	↗ ↗	↗ →	↗ –	↗ ↗	9; 14
Medium	19–46; 33.5/36.9	HMD vs Traditional method	↗ →	↗ –	↗ →	↗ ↗	9; 12
Medium	24,2	Smartphone vs Traditional method	↗	→	–	↗	14
Low – No Experience	30.5/40.9; 19–43; 19–40	Tablet vs Traditional method	– → ↘	↗ ↗ ↗	↗ – →	↗ – –	5; 26; 47
Low – No Experience	22–30; 29	Smartphone vs Traditional method	↗ ↘	→ →	– –	– –	15; 48
Low – No Experience	16–26; 19–43; 19–29; 23,77; 24,8; 21–46; 29; 22–52	HMD vs Traditional method	↘ → ↘ ↘ ↗ ↘ ↘ ↘	↗ ↗ ↗ ↗ ↗ ↘ → ↗	– – ↗ ↘ – – – –	↗ – – ↘ – – – –	18; 26; 30; 31; 42; 44; 48; 55
Low – No Experience	24,6; 19–37; 23–41; 21–46;	Projector vs Traditional method	↗ ↘ ↗ →	↗ → ↗ ↘	– ↗ – –	– – – –	20; 24; 40: 44

knowledge of AR. This changed when the operator had limited familiarity with the AR device; in fact, there was a worsening in task completion time, regardless of the age of the involved operators because the operator had no practicality with the device and took much more time to perform the task. Therefore, the same impact was the same if we considered operators from 19 to 29 and from 19 to 52 years of age. For the error rate, the results did not change; thanks to the AR device, the operator committed fewer errors than with traditional methods, even when the experience with AR was limited. This was linked to the fact that every action of the operator was monitored by the system, which could immediately detect incorrect actions and make corrections. The error correction procedure was carried out through a guided procedure. Conversely, in a classic manual assembly station, in the event of an operator-recognised error, it would take some time before the error was identified and consequently resolved, resulting in a greater number of errors being committed.

As reported in Table 7, regarding perceived workload and usability, the case studies produced opposite results; therefore, we could not determine whether there was a positive or a negative result. The variability of the results depended on the type of smart glasses used: some models generated low-quality images and involved a greater perceived workload by the user or HMD with low brightness, which, when in intense light, did not allow the information to be displayed effectively.

Concerning the comparison of 'Tablet vs Traditional method', the selected literature provided only case studies whose participants had a medium and low level of experience (Table 8). When the operator had a medium level of familiarity with AR and was aged 21 to 62, the operator performance using the tablet improved in terms of task completion time, perceived workload and usability. Concerning the error rate, the result obtained was not completely positive because there was a case study in which the use of a tablet was not significant compared with traditional methods. When the operator had little experience with the tablet and was aged 19 to 43, there was a slight worsening in terms of task completion time, while the result was improved in terms of error rate. Regarding perceived workload, the results were not completely positive.

Finally, the 'Smartphone vs Traditional method' comparison and results are shown in Table 9. When the operator had an average level of experience with

Table 7
Results of the "HMD vs Traditional method" comparison for each level of experience

HMD vs Traditional method	Metrics			
Experience with AR	Task completion time	Error rate	Perceived workload	Usability
High	↗	↗	↗	↗
Medium	↗	↗	↗	↗
Low	↘	↗	↗ / ↘	↗ / ↘

Table 8
Results of the "Tablet vs Traditional method" comparison for each level of experience

Tablet vs Traditional method	Metrics		
Experience with AR	Task completion time	Error rate	Perceived workload
Medium	↗	↗ / →	↗
Low	↘ / →	↗	↗ / →

Table 9
Results of the "Smartphone vs Traditional method" comparison for each level of experience

Smartphone vs Traditional method	Metrics	
Experience with AR	Task completion time	Error rate
Medium	↗	→
Low	↗ / ↘	→

the smartphone and average age of 24.2 years, there was an improvement in task completion time, while at the level of error rate, the use of the AR device was not significant compared with the traditional method because the number of errors was the same. Analysing operator performance with a low level of experience and aged 22 to 30, it could not be determined whether there was an improvement in task completion time because only two case studies were identified in the literature. One case study reported a positive result, and the other a negative result, but both agreed that the use of smartphones was not significant in terms of error rate reduction, regardless of the experience and age of the involved operators because in this case, it was also the same with respect to the traditional methods.

Discussions and conclusions

The growing interest of manufacturing companies in using digital technologies to improve the effectiveness and efficiency of production processes was analysed in this study which focused on AR solutions for improving human performance. To analyse the effects and benefits of AR devices on operator performance, objective and subjective metrics were evaluated for each AR devices.

In respect to the RQ1, the minimisation of the task completion time and the reduction in the error rate were considered objective metrics. The results showed that the shortest time to complete a task was achieved using AR devices (projectors, monitors, tablets or smartphones). For example, the use of the pick-by-projector reduced pick-up times by more than 70%, compared with when operators had to choose the correct component based on the information on the container label. However, the use of the HMD was disadvantageous because it increased the time to complete the activity, which was due to the participants' unfamiliarity with the device. Regarding the error rate, any AR device was better than traditional methods. The nature of the error depended on both the task performed and the circumstances of the working environment. In the case of manual assembly, some problems that could arise were an incorrect design of the components to be assembled or problems related to the level of information available, i.e., there were cases in which the information available to the operator was not accurate. In those cases, the operator, based on their experience, had to manage the information available pertaining to the actions to be performed.

With respect to subjective metrics, such as perceived workload, usability, learning rate and qualitative assessments, the analysis indicated that operators perceived that they had a lower workload with the use of the HMD or the monitor compared with traditional methods, while the use of the projector or tablet did not result in a perception of a lower workload. As for usability, scores on the SUS test were > 68, and this meant that the use of AR devices, particularly HMDs, projectors and tablets, was advantageous, although the users were more accustomed to paper manuals. The use of HMDs was not significant compared with traditional methods in terms of learning rate, which meant that the operator learned the information at the same speed, regardless of the tool used. Furthermore, the surveys and interviews of the operators indicated that they held a positive opinion related to projectors and tablets; however, they expressed negative feedback concerning HMDs. Operators complained about smart glasses because of their weight and bulk; also, they stated that after prolonged use, the glasses can cause headaches and dizziness. On the other hand, projectors were especially preferred in the field of training because they project information directly in the workplace without the operator having to look for instructions on how to perform a specific activity in paper manuals. The low cost of the tablet and the limited need for an experimentation phase were appreciated by operators. Senior operators who were already familiar with this device stated that the information could be viewed immediately, which was an extremely helpful feature.

For the RQ2, the age and experience of workers were also analysed. Twenty-four case studies involving operators under the age of 35, eight of which recruited operators aged 36 to 50, were found; no case studies were identified with workers over 50. This was a critical aspect of the study because AR should be tested to understand the effects on senior operators. Indeed, their cognitive, perceptual and motor skills can decline, and senior operators may suffer from a wide range of deficits, such as impaired vision and hearing, decreased working memory and decreased information processing speed. All of these problems could be solved by introducing AR within the company to assist the elderly employees. Finally, 14 case studies involving participants of different ages were identified; therefore, it was not possible to make a clear distinction between young and senior operators. Regarding the experience of the operators recruited in the selected case studies, the relationship of the operator with AR in 40 studies was mentioned, but upon further analysis, only two case studies in which the involved operators had a high experience with AR were identified. In three case studies, the level of knowledge of AR was medium, while in 25 case studies, the operators stated that they had little to no knowledge of AR. Based on the information collected and analysed, it was possible to evaluate in detail how the impacts on work performance vary with the age, experience and type of AR device, to understand if the variables 'age' and 'experience' play a crucial role with AR (Table 6). It is significant to note that the variable 'age' is an irrelevant parameter, if we consider operators aged under 50, while the level of 'experience' has a positive impact on performance; therefore, older operators could play a relevant role with the new technologies. Moreover, regardless of age and experience with any AR device, the operator could be completely unaware of the procedure that needs to be performed, but they could still be able to complete the task because this technology would guide them step-by-step. Therefore, an AR solution could reduce the amount of information required by the operator compared with that required by traditional methods.

Research gaps and future trends

The analysis showed that case studies on AR are still rather limited. Consequently, other case studies are needed to derive general conclusions that are statistically valid.

Concerning the attribute age, the conducted analysis leads to the conclusion that, in the analysed sample of case studies, the age is irrelevant, considering only case studies in which the operators had a low level of experience.

However, one research gap was identified through this study with respect to the 'age' of the involved operators, as no case study analysing the impacts of AR only on the performance of operators over 50 was identified. This is a critical aspect that deserves attention because AR technology could help the companies alleviate all sensory and cognitive effects on the operator as a result of advancing age; at the same time, 'experience' appears to have a positive impact on performance and, therefore, should be considered more for future case study research. So, first it should be investigated whether the 'age' of the operator can continue to be irrelevant (even if we are dealing with older operators), and if, a medium or high levels of experience, could affect the performance.

A second research gap revealed by this study relates to the type of comparison performed in the selected papers: they mainly analysed the parallelism between an AR device (in particular, HMD) and the traditional method; thus, the studies evaluating the comparison among multiple AR devices are limited. No case study was identified relating to the comparison of 'tablet vs monitor' or 'smartphone vs monitor'. For all the other combinations of technological devices, only one paper was found; hence, its information was insufficient to make a comparison and derive general conclusions.

Finally, with respect to the second analysis performed with the aim of identifying a relationship between AR devices and the age and experience of the operators, the following gaps were identified: (1) For those with a high level of experience, there were no comparisons between AR devices, such as 'Tablet vs Traditional method', 'Smartphone vs Traditional method' and 'Projector vs Traditional method'; (2) for those with a medium level of experience, there was only one comparison between AR devices, namely 'Smartphone vs Tablet'; and (3) for those with a low level experience, there were only three comparisons between AR devices, i.e., two related to 'Projector vs Monitor' and one related to 'HMD vs Tablet'.

It should also be considered that most of the case studies analysed were conducted in a laboratory environment and not in real production contexts. This aspect can limit the assessment of impacts on operator performance. Aspects such as interference with equipment or machinery, possible distractions linked to the use of AR technologies, analysis of the impact of the technologies with respect to the work shift (used for the entire shift or only in certain moments, time limit of use due to physical fatigue and/or or cognitive) have not been considered and require further investigation.

These aspects point out the necessity of providing new case studies in the literature to shed light on the benefits of AR and facilitate its integration at the industrial level.

AR technologies should be developed and tested taking into account the different objective attributes of the operators, such as age and experience with the technology, and above all with a greater focus on experience with the task. In particular, evaluating experience with the task could also help to define appropriate learning curves for operators engaged in complex and repetitive or non-repetitive tasks, allowing the development of adaptive and dynamic applications depending on the operators' learning level. Furthermore, decision support systems should be defined for choosing the most suitable technology based on the task and the attributes of the operators.

References

Alves, J.B. Marques, B., Dias, P., Sousa Santos, B. (2021). Using augmented reality for industrial quality assurance: a shop floor user study. *International Journal of Advanced Manufacturing Technology*, 115(1–2), 105–116.

Aschauer, A., Reisner-Kollmann, I., & Wolfartsberger, J. (2021). Creating an Open-Source Augmented Reality Remote Support Tool for Industry: Challenges and Learnings. *Procedia Computer Science*, 180, 269–279.

Aschenbrenner, D., Rojkov, M., Leutert, F. (2018). Comparing Different Augmented Reality Support Applications for Cooperative Repair of an Industrial Robot. *Adjunct Proceedings – 2018 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct*, 69–74.

Barbieri, L. & Marino, E. (2019). An Augmented Reality Tool to Detect Design Discrepancies: A Comparison Test with Traditional Methods. Lecture Notes in Computer Science.Springer International Publishing.

Blattgerste, J. Strenge, B., Renner, P., Pfeiffer, T., Essig, K. (2017). Comparing conventional and augmented reality instructions for manual assembly tasks. *ACM International Conference Proceeding Series.* Part F1285, 75–82.

Bonavolontá, F., Dallet, D., Erra, E., Grassi, A., Popolo, V., Tedesco, A., & Vespoli, S. (2020). Measuring Worker' s Performance in Augmented Reality-assisted Industry 4.0 Procedures. *IEEE International Instrumentation and Measurement Technology Conference (I2MTC).* 1–6.

Bosch, T., Könemann, R., De Cock, H., & Van Rhijn, G. (2017). The effects of projected versus display instructions on productivity, quality and workload in a simulated assembly task. *ACM International Conference Proceeding Series.* Island of Rhodes, Greece, 412–415.

Bosch, T., Van Rhijn, G., Krause, F., Könemann, R., Wilschut, E.S., & de Looze, M. (2020). Spatial augmented reality: A tool for operator guidance and training evaluated in five industrial case studies. *ACM International Conference Proceeding Series.* 296–302.

Bottani, E., Longo, F., Nicoletti, L., Padovano, A., Tancredi, G.P.C., Tebaldi, L., Vetrano, M., & Vignali, G. (2021). Wearable and interactive mixed reality solutions for fault diagnosis and assistance in manufacturing systems: Implementation and testing in an aseptic bottling line. *Computers in Industry*, 128, 103429.

Bottani, E. & Vignali, G. (2019). Augmented reality technology in the manufacturing industry: A review of the last decade. *IISE Transactions*, 51(3), 284–310.

Brice, D., Rafferty, K. and McLoone, S. (2020). AugmenTech: The Usability Evaluation of an AR System for Maintenance in Industry. Lecture Notes in Computer Science, 12243 LNCS, 284–303.

Brizzi, F., Peppoloni, L., Graziano, A., Di Stefano, E., Avizzano, C. A., & Ruffaldi, E. (2018). Effects of Augmented Reality on the Performance of Teleoperated Industrial Assembly Tasks in a Robotic Embodiment. *IEEE Transactions on Human-Machine Systems*, 48(2), 197–206.

Bruno, F., Barbieri, L., Marino, E., Muzzupappa, M., D'Oriano, L., & Colacino, B. (2019). An augmented reality tool to detect and annotate design variations in an Industry 4.0 approach. *International Journal of Advanced Manufacturing Technology*, 105(1–4), 875–887.

Büttner, S., Funk, M., Sand, O., & Röcker, C. (2016). Using head-mounted displays and in-situ projection for assistive systems – A comparison. *ACM International Conference Proceeding Series*, 29-June-20.

Büttner, S., Prilla, M., & Röcker, C. (2020). Augmented Reality Training for Industrial Assembly Work – Are Projection-based AR Assistive Systems an Appropriate Tool for Assembly Training? Conference on Human Factors in Computing Systems – Proceedings, 1–12.

Danielsson, O., Holm, M., & Syberfeldt, A. (2020). Augmented reality smart glasses in industrial assembly: Current status and future challenges. *Journal of Industrial Information Integration*, 20, 100175.

Di Pasquale, V. De Simone, V., Miranda, S., & Riemma, S. (2022). Smart Operators: How Augmented and Virtual Technologies Are Affecting the Worker's Performance in Manufacturing Contexts. *Journal of Industrial Engineering and Management*, 15(2), 233–235.

Doshi, A., Smith, R.T., Thomas, B.H., & Bouras, C. (2017). Use of projector based augmented reality to improve manual spot-welding precision and accuracy for automotive manufacturing. *International Journal of Advanced Manufacturing Technology*, 89(5–8), 1279–1293.

Fang, W. & An, Z. (2020). A scalable wearable AR system for manual order picking based on warehouse floor-related navigation. *International Journal of Advanced Manufacturing Technology*, 109(7–8), 2023–2037.

Fiorentino, M., Radkowski, R., Boccaccio, A., & Uva, A. E. (2016). Magic mirror interface for augmented reality maintenance: An automotive case study. *Proceedings of the Workshop on Advanced Visual Interfaces AVI*, 07-10-June-2016, 160–167.

Funk, M., Shirazi, A.S., Mayer, S., Lischke, L., & Schmidt, A. (2015). Pick from Here-An interactive mobile cart using in-situ projection for order picking. *UbiComp 2015 – Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 601–609.

Funk, M., Mayer, S., Nistor, M., & Schmidt, A. (2016). Mobile In-Situ Pick-by-Vision: Order Picking Support using a Projector Helmet. *PETRA '16: Proceedings of the 9th ACM International Conference on PErvasive Technologies Related to Assistive Environments,* 1–4.

Funk, M., Bächler, A., Bächler, L., Kosch, T., Heidenreich, T., & Schmidt, A. (2017). Working with augmented reality? A long-term analysis of in-situ instructions at the assembly workplace. *ACM International Conference Proceeding Series*, Part F128530, 222–229.

Funk, M., Kosch, T. & Schmidt, A. (2016). Interactive worker assistance: Comparing the effects of in-situ projection, head-mounted displays, tablet, and paper instructions. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 934–939.

Gattullo, M., Evangelista, A., Uva, A.E., Fiorentino, M., Boccaccio, A., & Manghisi, V. M. (2019). Exploiting Augmented Reality to Enhance Piping and Instrumentation Diagrams for Information Retrieval Tasks in Industry 4.0 Maintenance. *Lecture Notes in Computer Science*. Springer International Publishing.

Havard, V., Baudry, D., Savatier, X., Jeanne, B., Louis, A., & Mazari, B. (2016). Augmented industrial maintenance (AIM): A case study for evaluating and comparing with paper and video media supports. *Lecture Notes in Computer Science*, 9768, 302–320.

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(4), 999–1014.

Hietanen, A., Pieters, R., Lanz, M., Latokartano, J., & Kämäräinen, J. K. (2020). AR-based interaction for human-robot collaborative manufacturing. *Robotics and Computer-Integrated Manufacturing*, 63(2019), 101891.

Ikiz, Y.D., Atici-Ulusu, H., Taskapilioglu, O., Gunduz, T. (2019). Effects of Augmented Reality Glasses on the Cognitive Load of Different Age Groups. *International conference on Modern Research in Science, Engineering and Technology*, 20–24.

Illing, J., Klinke, P., Grünefeld, U., Pfingsthorn, M., & Heuten, W. (2020). Time is money! Evaluating Augmented Reality Instructions for Time-Critical Assembly Tasks. *ACM International Conference Proceeding Series*, 277–287.

Konstantinidis, F.K., Kansizoglou, I., Santavas, N., Mouroutsos, S.G., & Gasteratos, A. (2020). Marma: A mobile augmented reality maintenance assistant for fast-track repair procedures in the context of industry 4.0. *Machines*, 8(4), 1–15.

Koumaditis, K., Venckute, S., Jensen, F.S., & Chinello, F. (2019). Immersive Training: Outcomes from Small Scale AR/VR Pilot-Studies. *26th IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2019 – Proceedings*, 1894–1898.

Kubenke, J. and Kunz, A. (2019). Potentials of IT-supported assistive systems: Comparison of two user studies in the manufacturing industry. *IFAC-PapersOnLine*, 52(13), 1866–1871.

Lagorio, A. et al. (2022). Augmented Reality in Logistics 4.0: implications for the human work. *IFAC-PapersOnLine*, 1–6.

Lai, Z.H., Tao, W., Leu, M.C., & Yin, Z. (2020). Smart augmented reality instructional system for mechanical assembly towards worker-centered intelligent manufacturing. Journal of Manufacturing Systems, 55, 69–81.

Lampen, E., Teuber, J., Gaisbauer, F., Bär, T., Pfeiffer, T., & Wachsmuth, S. (2019). Combining simulation and augmented reality methods for enhanced worker assistance in manual assembly. *Procedia CIRP*, 81(July), 588–659.

Leutert, F. & Schilling, K. (2018). Projector-based Augmented Reality for Telemaintenance Support. *IFAC-PapersOnLine*, 51(11), 502–507.

Loch, F., Quint, F., & Brishtel, I. (2016). Comparing video and augmented reality assistance in manual assembly. *Proceedings – 12th International Conference on Intelligent Environments, IE 2016*, 147–150.

Longo, F., Nicoletti, L., & Padovano, A. (2017). Smart operators in industry 4.0: A human-centered approach to enhance operators' capabilities and competencies within the new smart factory context. *Computers and Industrial Engineering*, 113, 144–159.

Marino, E., Barbieri, L., Colacino, B., Fleri, A.K., & Bruno, F. (2021). An Augmented Reality inspection tool to support workers in Industry 4.0 environments. *Computers in Industry*, 127, 103412.

Mättig, B. & Kretschmer, V. (2019). Smart packaging in intralogistics: An evaluation study of human-technology interaction in applying new collaboration technologies. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 739–748.

Mengoni, M., Ceccacci, S., Generosi, A., & Leopardi, A. (2018). Spatial Augmented Reality: An application for human work in smart manufacturing environment. *Procedia Manufacturing*, 17, 476–483.

Moghaddam, M., Wilson, N.C., Modestino, A.S., Jona, K., & Marsella, S.C. (2021). Exploring augmented reality for worker assistance versus training. *Advanced Engineering Informatics*, 50, 101410.

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. *International Journal of Advanced Manufacturing Technology*, 105(9), 3899–3910.

Murauer, N., Panz, N., & Von Hassel, C. (2018). Comparison of scan mechanisms in augmented reality supported order picking processes. *CEUR Workshop Proceedings*, 2082, 69–76.

Palmarini, R., Erkoyuncu, J.A., Roy, R., & Torabmostaedi, H. (2018). A systematic review of augmented reality applications in maintenance. *Robotics and Computer-Integrated Manufacturing,* 49, 215–228.

Pilati, F., Faccio, M., Gamberi, M., & Regattieri, A. (2020). Learning manual assembly through real-time motion capture for operator training with augmented reality. *Procedia Manufacturing*, 45, 189–195.

Princle, A., Campbell, A.G., Hutka, S., Torrasso, A. (2018). Using an Industry-Ready AR HMD on a Real Maintenance Task: AR Benefits Performance on Certain Task Steps More Than Others. *Adjunct Proceedings – 2018 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2018*, 236–241.

Ramírez, H., Mendoza, E., Mendoza, M., & González, E. (2015). Application of Augmented Reality in Statistical Process Control, to Increment the Productivity in Manufacture. *Procedia Computer Science*, 75, 213–220.

Re, G. M., Oliver, J., & Bordegoni, M. (2016). Impact of monitor-based augmented reality for on-site industrial manual operations. *Cognition, Technology and Work*, 18(2), 379–392.

Rice, M., Ma, K.T., Tay, H.H., Kaliappan, J., Koh, W.L., Tan, W.P., & Ng, J. (2018). Evaluating an augmented remote assistance platform to support industrial applications. *IEEE World Forum on Internet of Things, WF-IoT 2018 – Proceedings*, 592–597.

Romero, D., Stahre, J., Wuest, T., Noran, O., Bernus, P., Fast-Berglund, Å., & Gorecky, D. (2016). Towards an operator 4.0 typology: A human-centric perspective on the fourth industrial revolution technologies. *CIE 2016: 46th International Conferences on Computers and Industrial Engineering, (October)*, 1–11.

Romero, D., Stahre, J., & Taisch, M. (2020). The Operator 4.0: Towards socially sustainable factories of the future. *Computers and Industrial Engineering*, 139(November 2019).

Rupprecht, P., Kueffner-McCauley, H., & Schlund, S. (2020). Information provision utilizing a dynamic projection system in industrial site assembly. *Procedia CIRP*, 93, 1182–1187.

Saeed, N. (2021). Implementing the Augmented Reality as an Industry 4.0 Application to Simplify the Busbar Bending Process during the Covid-19 Pandemic. *Transactions of FAMENA*, 115–125.

Schuster, F., Engelmann, B., Sponholz, U., Schmitt, J., & Engineering, I.D. (2021). Human acceptance evaluation of AR-assisted assembly scenarios. *Journal of Manufacturing Systems*, (April 2020).

Sharma, A., Mehtab, R., Mohan, S., & Mohd Shah, M.K. (2022). Augmented reality – an important aspect of Industry 4.0. *Industrial Robot*, 49(3), 428–441.

Syberfeldt, A., Danielsson, O., Holm, M., & Wang, L. (2015). Visual Assembling Guidance Using Augmented Reality. *Procedia Manufacturing*, 1, 98–109.

Wang, C.H., Tsai, N.H., Lu, J.M., & Wang, M.J.J. (2019). Usability evaluation of an instructional application based on Google Glass for mobile phone disassembly tasks. *Applied Ergonomics*, 77, 58–69.

Werrlich, S., Lorber, C., Nguyen, P.A., Yanez, C.E.F., & Notni, G. (2018). Assembly training: Comparing the effects of head-mounted displays and face-to-face training. *Lecture Notes in Computer Science*. Springer International Publishing.

Werrlich, S., Nguyen, P.A., & Notni, G. (2018). Evaluating the training transfer of Head-Mounted Display based training for assembly tasks. *ACM International Conference Proceeding Series*, 297–302.

Wilschut, E.S., Murphy, M.S. & Bosch, T. (2019). Evaluating learning approaches for product assembly: using chunking of instructions, spatial augmented reality and display based work instructions. *Petra'19*, 376–381.

Yang, Y., Karreman, J. & De Jong, M. (2020). Comparing the Effects of Paper and Mobile Augmented Reality Instructions to Guide Assembly Tasks. *IEEE International Professional Communication Conference*, 96–104.

APPENDIX A1

ID	References	Devices	Type of application	Type of task	Age	Experience (with AR and/or with task)
1	(Havard et al., 2021)	Tablet, Paper	Assistance	Maintenance	[20-25]	WITH TASK: Beginner, Intermediate and Advanced operators
2	(Moghaddam et al., 2021)	HMD, Paper	Assistance +Training	Assembly	n.a.	WITH AR: No or minor prior experience
3	(Alves et al., 2021)	Monitor, Paper	Assistance	Quality control	n.a.	WITH TASK: 4 operators had never performed the quality control test before (inexperienced users) while 3 perform this task routinely
4	(Bottani et al., 2021)	HMD, Smartphone	Assistance+ Training	Health and Safety	[31-51]	WITH AR: 3 of 8 operators are expected to handle and make use of the devices
5	(Marino et al., 2021)	Tablet, Paper	Assistance	Health and Safety	G1: Avg. age 30,5 G2: Avg. age 40,9	WITH AR: All of the operators were novices to AR technology
6	(Dila IKIZ et al., 2019)	HMD, Paper	Assistance	Assembly	Under 35 / Over 35	–
7	(Aschauer et al., 2021)	Smartphone, Paper	Remote collaboration	Maintenance	n.a.	WITH AR and TASK: Different technical skill levels
8	(Konstantinidis et al., 2020)	Smartphone, Paper	Training	Maintenance	n.a.	
9	(Illing et al., 2020)	HMD, Tablet, Paper	Training	Assembly	Tablet vs Paper: [21-62] HMD vs Paper: [19-46]	WITH AR: The participants had limited to medium experience with AR WITH TASK: They had medium to very high experience with work instructions
10	(Bosch et al., 2020)	Projector, Paper	Assistance +Training	Assembly	n.a.	WITH AR and TASK: Experienced operators (paper) vs novice operator (AR)
11	(Lai et al., 2020)	Monitor, Paper	Assistance	Assembly	Avg. age 28	WITH TASK: No one had a prior knowledge of the experiment
12	(Brice et al., 2020)	HMD, Paper	Assistance	Maintenance	Avg. age 33,5 and 36,9	WITH AR: Many of the participants had previously tried the HoloLens at demonstrations around the university.
13	(Bruno et al., 2019)	Group 1 and Group 2	Remote collaboration	Design variations	G1: Avg. age 48,5; G2: Avg. age 36,8	WITH AR: Both groups of participants were naive with AR technologies.
14	(Gattullo et al., 2019)	Smartphone, Tablet, Paper	Training	Maintenance	Avg age 24,2	WITH AR: For as the frequency of usage of AR applications: 13 never, 15 rarely, 8 sometimes, 3 often, and none always used AR applications

ID	References	Devices	Type of application	Type of task	Age	Experience (with AR and/or with task)
15	(Barbieri and Marino, 2019)	Smartphone, Paper	Remote collaboration	Design variations	[22-30]	WITH AR: None of the participants had any previous experience WITH TASK: All of them were familiar with 2D engineering drawings and expert in the use of 3D modelling and CAD systems
16	(Lampen et al., 2019)	HMD, Paper	Assistance	Assembly	[18-54]	WITH AR: No one had a considerable experience with AR methods
17	(Werrlich, Nguyen and Notni, 2018a)	Group 1 and Group 2	Training	Assembly	G1: (G2: (	WITH TASK: Group 1 and Group 2 stated to have medium experience with assembly processes
18	(Werrlich et al., 2018b)	HMD, Paper	Training	Assembly	[16-26]	WITH AR: All participants had no prior experience in using or working with AR
19	(Murauer et al., 2018)	Monitor (M), HMD (AR 1), HMD+ scan glove (AR 2)	Assistance	Order Picking	Avg. age 25	–
20	(Werrlich et al., 2018b)	Projector, Paper	Assistance+ Training	Maintenance	Avg. age 24,6	WITH AR: 14 participants were novices to AR and computer-based maintenance guidance system WITH TASK: Half of the participants had previous experience with maintenance manual instructions
21	(Bosch et al., 2017)	Projector, Monitor	Assistance+ Training	Assembly	Avg. age 40,2	WITH TASK: Experienced (i.e., assembly workers) and unexperienced workers (i.e., engineers and researchers)
22	(Funk et al., 2017)	Projector, Paper	Assistance+ Training	Assembly	Avg. age 43,34 and 45,67	WITH AR and TASK: Expert workers/ Untrained workers
23	(Doshi et al., 2017)	Projector, Paper	Assistance	Quality control	n.a	WITH AR and TASK: Highly skilled automotive welders
24	(Funk et al., 2016a)	HelmetPickAR, Pick by-Paper	Assistance	Order Picking	[19-37]	WITH AR: All participants were using HelmetPickAR for the first time WITH TASK: None of the participants were familiar with the order picking
25	(Re et al., 2016)	Monitor, Paper	Assistance	Assembly	n.a.	–
26	(Havard et al., 2016)	AR media (HMD and Tablet), Paper	Remote collaboration	Maintenance	[19-43]	WITH AR: No experience
27	(Funk et al., 2015)	OrderPickAR (projector), Pick by Paper, Pick by Vision (HMD)	Assistance	Order Picking	n.a.	–

ID	References	Devices	Type of application	Type of task	Age	Experience (with AR and/or with task)
28	(Ramírez et al., 2015)	Tablet, Paper	Training	Quality control	n.a.	WITH AR: The participants had a basic expertise on mobile devices
29	(Fang and An, 2020)	Pick by Paper, Pick by Vision (HMD)	Assistance	Order Picking	n.a.	WITH AR and TASK: Workers without experience
30	(Mourtzis, Zogopoulos and Xanthi, 2019)	HMD, Paper	Assistance	Assembly	[19-29	WITH AR: No one from the participants had ever used Augmented Reality for instructions visualization before WITH TASK: The operators had limited previous experience in assembly tasks, and none of them had not performed this assembly task before
31	(Wang et al., 2019)	HMD, Paper	Assistance	Maintenance	Avg. age 23,77	WITH AR: The recruited participants had no prior experience using an HMD WITH TASK: They had no prior experience about disassembling a smart device
32	(Koumaditis et al., 2019)	HMD, Paper	Training	Assembly	n.a.	WITH AR: No experience with AR (trainee)
33	(Mengoni et al., 2018)	Projector, Monitor	Assistance +Training	Assembly	[21-27]	WITH AR: The participants were not familiar with Samsung Smartphones
34	(Saeed, 2021)	Smartphone, Paper	Training	Quality control	Avg. age 32	WITH TASK: No experience on the bending machine
35	(Pilati et al., 2020)	Monitor and Paper	Assistance+ Training	Assembly	n.a	WITH AR and TASK: Inexperienced operators
36	(Wilschut et al., 2019)	Projector, Paper	Training	Assembly	Avg. age 24	WITH AR and TASK: Inexperienced operators
37	(Mättig and Kretschmer, 2019)	HMD, Monitor, Paper	Assistance	Packaging	[20-40]	WITH AR: The subjects had a high acceptance towards technologies and a relatively high level of the need using technology
38	(Aschenbrenner et al., 2018)	Projector and Tablet, HMD, Smartphone	Remote collaboration	Maintenance	Avg. age 21,1	
39	(Brizzi et al., 2018)	HMD, Paper	Remote collaboration	Pick and place	[23-40]	WITH AR: Experts in AR/VR
40	(Fiorentino et al., 2016)	Projector, Paper	Training	Maintenance	[23-41]	WITH AR: for laboratory 2 of the participants used virtual reality systems; for industry none of them used virtual reality systems before
41	(Hietanen et al., 2020)	HMD, Projector, Paper	Assistance	Assembly	n.a.	WITH AR and TASK: No experience

ID	References	Devices	Type of application	Type of task	Age	Experience (with AR and/or with task)
42	(Bonavolontŕ et al., 2020)	HMD, Paper	Training	Assembly	Avg. age 24,8	WITH AR: None of the participants was familiar with AR
43	(Loch et al., 2016)	Monitor, Paper	Assistance+ Training	Assembly	n.a.	WITH AR: None of the participants had previous experience with AR-based assistance systems
44	(Büttner et al., 2016)	HMD, Projector, Paper	Assistance+ Training	Assembly	[21-46]	WITH AR: None of the participants had worked with AR
45	(Büttner et al., 2020)	Projector, Paper	Training	Assembly	n.a.	–
46	(Blattgerste et al., 2017)	Smartphone, HMD, Paper	Assistance+ Training	Assembly	[20-33]	WITH TASK: They had no prior experience with assembly task
47	(Yang et al., 2020)	Tablet, Paper	Assistance	Assembly	[19-40]	WITH AR: Only five participants reported that they had experience with AR instructions
48	(Rice et al., 2018)	HMD, Smartphone, Paper	Remote collaboration	Assembly	Avg. age 29	WITH AR: None of the participants were involved in the previous study with no known experience
49	(Schuster et al., 2021)	HMD and Without HMD	Assistance	Assembly	Avg. age 28,5	WITH TASK: None of the participants has performed a similar task before
50	(Rupprecht et al., 2020)	Projector, Monitor	Assistance	Assembly	12 participants [20-29] and 4 [30-39]	WITH AR: On average 11 people had no experience with augmented reality
51	(Funk, et al., 2016b)	HMD, Tablet, Projector, Paper	Assistance+ Training	Assembly	between 20-33	WITH TASK: Participants were not familiar with assembly task
52	(Kubenke and Kunz, 2019)	HMD, Tablet, Paper	Assistance	Assembly	n.a.	–
53	(Leutert and Schilling, 2018)	HMD, projector, Tablet, Paper	Remote collaboration	Maintenance	n.a.	–
54	(Princle et al., 2018)	HMD, Paper	Training	Maintenance	Avg. age 22,44	WITH AR: All participants had never previously used the AR Yaw brake service application and only one had previously used the HMD
55	(Syberfeldt et al., 2015)	HMD, Paper	Training	Assembly	[22-52]	WITH AR: None of the subjects had any previous experience of AR WITH TASK: None of the subjects had any previous experience of industrial manufacturing or assembling

APPENDIX A2

	Technology		Objective metrics		Subjective metrics			
ID	Device 1	Device 2	Task Completion Time	Error Rate	Perceived workload	Usability	Learning rate	Qualitative evaluation
1	Tablet	Paper	$\nearrow$ Tablet $<$ Paper	$\nearrow$ Tablet $<$ Paper	$\rightarrow$ Tablet $\approx$ Paper	$\searrow$ Tablet $<$ Paper	-	-
1	Tablet	Paper	$\nearrow$ Tablet $<$ Paper	$\nearrow$ Tablet $<$ Paper	$\rightarrow$ Tablet $\approx$ Paper	$\rightarrow$ Tablet $\approx$ Paper	-	-
2	HMD	Paper	$\searrow$ HMD $>$ Paper	$\nearrow$ HMD $<$ Paper	$\rightarrow$ HMD $\approx$ Paper	-	$\nearrow$ HMD $>$ Paper	-
3	Monitor	Paper	$\nearrow$ Monitor $<$ Paper			–	–	–
4	HMD	Smartphone				$\searrow$ HMD $<$ Smartphone	$\nearrow$ HMD $>$ Smartphone	–
5	Tablet	Paper		$\nearrow$ Tablet $<$ Paper	$\nearrow$ Tablet $<$ Paper	$\nearrow$ Tablet $<$ Paper	-	-
6	HMD	Paper			$\nearrow$ HMD $<$ Paper	–	–	–
6	HMD	Paper			$\nearrow$ HMD $<$ Paper	–	–	–
7	Smartphone	Paper	$\rightarrow$ Smartphone $\approx$ Paper	$\nearrow$ Smartphone $<$ Paper				
8	Smartphone	Paper	$\nearrow$ Smartphone $<$ Paper					
9	Tablet	Paper	$\nearrow$ Tablet $<$ Paper	$\nearrow$ Tablet $<$ Paper	$\nearrow$ Tablet $<$ Paper	$\nearrow$ Tablet $>$ Paper		$\nearrow$ Tablet $>$ Paper
9	HMD	Paper	$\nearrow$ HMD $<$ Paper	$\nearrow$ HMD $<$ Paper	$\nearrow$ HMD $<$ Paper	$\nearrow$ HMD $>$ Paper		$\rightarrow$ HMD $\approx$ Paper
10	Projector	Paper	$\searrow$ Projector $>$ Paper	$\searrow$ Projector $>$ Paper	$\rightarrow$Projector $\approx$ Paper	$\nearrow$ Projector $>$ Paper		
10	Projector	Paper	$\searrow$ Projector $>$ Paper	$\rightarrow$ Projector $\approx$ Paper		$\nearrow$ Projector $>$ Paper	$\nearrow$ Projector $>$ Paper	
11	Monitor	Paper	$\nearrow$ Monitor $<$ Paper	$\nearrow$ Monitor $<$ Paper				
12	HMD	Paper	$\rightarrow$ HMD $\approx$ Paper		$\rightarrow$ HMD $\approx$ Paper	$\nearrow$ HMD $>$ Paper		
13	Smartphone		$\rightarrow$ G1 $\approx$ G2	$\rightarrow$ G1 $\approx$ G2				$\searrow$ G1 $<$ G2
14	Smartphone	Paper	$\nearrow$ Smartphone $<$ Paper	$\rightarrow$ Smartphone $\approx$ Paper		$\nearrow$ Smartphone $>$ Paper		

	Technology		Objective metrics		Subjective metrics			
ID	Device 1	Device 2	Task Completion Time	Error Rate	Perceived workload	Usability	Learning rate	Qualitative evaluation
14	Tablet	Paper	↗ Tablet < Paper	→ Tablet ≈ Paper		↗ Tablet > Paper		
14	Smartphone	Tablet	→ Smartphone ≈ Tablet	→ Smartphone ≈ Tablet		→ Smartphone ≈ Tablet		
15	Smartphone	Paper	↗ Smartphone < Paper	→ Smartphone ≈ Paper				
16	HMD	Paper	→ HMD ≈ Paper	↗ HMD < Paper	↗ HMD < Paper	↗ HMD < Paper	→ HMD ≈ Paper	
17	HMD			↘ G1 > G2	→ G1 ≈ G2	→ G1 ≈ G2		
18	HMD	Paper	↘ HMD > Paper	↗ HMD < Paper		↗ HMD < Paper	→ HMD ≈ Paper	
19	HMD (AR 1)	Monitor (M)	↘ AR 1 > M	↗ AR 1 < M	↘ AR 1 > M	↘ AR 1 < M		
19	HMD + scan glove (AR 2)	Monitor (M)	↗ AR 2 < M	↗ AR 2 < M	→ AR 2 ≈ M			
19	AR 2	AR 1	↗ AR 2 < AR 1	↘ AR 2 > AR 1	↗ AR 2 < AR 1	↗ AR 2 > AR 1		
20	Projector	Paper	↗ Projector < Paper	↗ Projector < Paper				↗ Projector > Paper
21	Projector	Monitor	↗ Projection < Monitor	↗ Projection < Monitor	↗ Projection < Monitor			
22	Projector	Paper	↘ Projector > Paper	→ Projector = Paper	→ Projector = Paper			
22	Projector	Paper	↘ Projector > Paper After 3 days: ↗ Projector < Paper	↘ Projector > Paper After 3 days: ↗ Projector < Paper	→ Projector = Paper			↗ Projector < Paper
23	Projector	Paper		↗ Projector < Paper				
24	Helmet PickAR (Projector)	Pick-by-Paper	Pick time: → Helmet-PickAR = Pick-by-Paper Place time: ↘ Helmet-PickAR > Pick-by-Paper	→ Helmet-PickAR = Pick-by-Paper	↗ Helmet-PickAR < Pick-by-Paper			↗ Helmet-PickAR > Pick-by-Paper

	Technology		Objective metrics		Subjective metrics			
ID	Device 1	Device 2	Task Completion Time	Error Rate	Perceived workload	Usability	Learning rate	Qualitative evaluation
25	Monitor	Paper	↗ Monitor < Paper	→ Monitor=Paper	↗ Monitor < Paper			↗ Monitor > Paper
26	HMD	Paper	→ HMD = Paper	↗ HMD < Paper				↗ HMD < Paper
26	Tablet	Paper	→ Tablet = Paper	↗ Tablet < Paper				↗ Tablet > Paper
27	OrderPickAR (projector)	Pick by Paper	↗ Order-PickAR < Pick by Paper	↗ Order-PickAR < Pick by Voice	↗ Order-PickAR < Pick by Paper			
27	OrderPickAR (projector)	Pick by Vision (HMD)	↗ Order-PickAR < Pick by Vision	↗ Order-PickAR < Pick by Vision	↗ Order-PickAR < Pick by Vision			
27	Pick by Vision (HMD)	Pick by Paper	↘ Pick by Vision > Paper	↘ Pick by Vision > Paper	↘ Pick by Vision > Paper			
28	Tablet	Paper	↗ Tablet < Paper					
29	Pick by AR (HMD)	Paper	↗ Pick by AR < Paper	↗ Pick by AR < Paper	↗ Pick by AR < Paper	↗ Pick by AR > Paper		
30	HMD	Paper	↘ HMD > Paper	↗ HMD < Paper	↗ HMD < Paper			
31	HMD	Paper	↘ HMD > Paper	↗ HMD < Paper	↘ HMD > Paper	↘ HMD < Paper		
32	HMD	Paper	↗ HMD < Paper		↗ HMD < Paper			
33	Projector	Monitor	↗ Projector < Monitor		→ Projector = Monitor			
34	Smartphone	Paper		↗ Smart-phone < Paper				
35	Monitor	Paper	↗ Monitor < Paper				↗ Monitor > Paper	
36	Projector	Paper	→ Projector = Paper	→ Projector = Paper	→ Projector = Paper			
37	HMD	Monitor		↗ HMD < Monitor	→ HMD = Monitor	↘ HMD < Monitor		↘ HMD < Monitor
37	HMD	Paper		↗ HMD < Paper	↗ HMD < Paper	↗ HMD > Paper		
37	Monitor	Paper		↗ Monitor < Paper	↗ Monitor < Paper	↗ Monitor > Paper		
38	Projector	HMD	↗ Projector < HMD		↗ Projector < HMD			

	Technology		Objective metrics		Subjective metrics			
ID	Device 1	Device 2	Task Completion Time	Error Rate	Perceived workload	Usability	Learning rate	Qualitative evaluation
38	Projector	Smartphone	↗ Projector < Smart-phone		↗ Projector < Smart-phone			
38	Projector	Tablet	↗ Projector < Tablet		↗ Projector < Tablet			
39	HMD	Paper	↗ HMD < Paper	↗ HMD < Paper				
40	Projector	Paper	↗ Projector < Paper	↗ Projector < Paper				
41	HMD	Projector	↘ HMD > Projector					↘ HMD < Projector
41	HMD	Paper	↗ HMD < Paper					↘ HMD < Paper
41	Projector	Paper	↗ Projector < Paper					↗ Projector > Paper
42	HMD	Paper	↗ HMD < Paper	↗ HMD < Paper				
43	Monitor	Paper	→ Monitor = Paper	↗ Monitor < Paper	↗ Monitor < Paper			
44	Projector	HMD	↗ Projector < HMD	↗ Projector < HMD				↗ Projector > HMD
44	Projector	Paper	→ Projector = Paper	↘ Projector > Paper				→ Projector = Paper
44	HMD	Paper	↘ HMD > Paper	↘ HMD > Paper				↘ HMD > Paper
45	Projector	Paper	Day 1: → Projector = Paper Day 2: ↗ Projector < Paper	After 24 hours: → Projector = Paper After one week: → Projector = Paper				
46	HMD (Epson Moverio)	Paper	→ HMD (Epson Moverio) = Paper	→ HMD (Epson Moverio) = Paper	↘ HMD (Epson Moverio) > Paper			↘ HMD (Epson Moverio) < Paper
46	HMD (HoloLens)	Paper	→ HMD (HoloLens) = Paper	→ HMD (HoloLens) = Paper	↘ HMD (HoloLens) > Paper			↘ HMD (HoloLens) < Paper

	Technology		Objective metrics		Subjective metrics			
ID	Device 1	Device 2	Task Completion Time	Error Rate	Perceived workload	Usability	Learning rate	Qualitative evaluation
46	HMD (Epson Moverio)	Smartphone	→ HMD (Epson Moverio) = Smartphone	→ HMD (Epson Moverio) = Smartphone	↗ HMD (Epson Moverio) < Smartphone			↗ HMD (Epson Moverio) > Smartphone
46	HMD (HoloLens)	Smartphone	→ HMD (HoloLens) = Smartphone	→ HMD (HoloLens) = Smartphone	→ HMD (HoloLens) = Smartphone			↗ HMD (HoloLens) > Smartphone
46	HMD (Epson Moverio)	HMD (HoloLens)	→ HMD (Epson Moverio) = HMD (HoloLens)	→ HMD (Epson Moverio) = HMD (HoloLens)	↗ HMD (Epson Moverio) < HMD (HoloLens)			↘ HMD (Epson Moverio) < HMD (HoloLens)
46	Smartphone	Paper	→ Smartphone = Paper	→ Smartphone = Paper	↘ Smartphone > Paper			↘ Smartphone < Paper
47	Tablet	Paper	↘ Tablet > Paper	↗ Tablet < Paper	→ Tablet = Paper			
48	HMD	Smartphone	→ HMD = Smartphone	→ HMD=Smartphone				
48	HMD	Paper	↘ HMD > Paper	→ HMD = Paper				
48	Smartphone	Paper	↘ Smartphone > Paper	→Smartphone=Paper				
49	HMD	Paper	↗ HMD < Paper					
50	Projector	Monitor	↗ Projector < Monitor			↗ Projector > Monitor		↗ Projector > Monitor
51	Projector	HMD	↗ Projector < HMD	↗ Projector < HMD	↗ Projector < HMD			
51	Projector	Tablet	→ Projector = Tablet	↗ Projector < Tablet	→ Projector = Tablet			
51	Projector	Paper	→ Projector = Paper	→ Projector = Paper	→ Projector = Paper			
51	HMD	Tablet	→ HMD = Tablet	→ HMD = Tablet	→ HMD = Tablet			
51	HMD	Paper	→ HMD = Paper	→ HMD = Paper	→ HMD = Paper			
51	Tablet	Paper	→ Tablet = Paper	→ Tablet = Paper	→ Tablet = Paper			
52	HMD	Paper	↗ HMD < Paper	↗ HMD < Paper	↘ HMD > Paper			

	Technology		Objective metrics		Subjective metrics			
ID	Device 1	Device 2	Task Completion Time	Error Rate	Perceived workload	Usability	Learning rate	Qualitative evaluation
52	HMD	Tablet	↗ HMD < Tablet	↗ HMD < Tablet	↘ HMD > Tablet			
52	Tablet	Paper			↘ Tablet > Paper			
53	Projector	HMD	↗ Projector < HMD	↗ Projector < HMD	↗ Projector < HMD			
53	Projector	Tablet	↗ Projector < Tablet	↗ Projector < Tablet	↗ Projector < Tablet			
53	Projector	Paper	↗ Projector < Paper	↗ Projector < Paper	↗ Projector < Paper			
54	HMD	Tablet	↗ HMD < Tablet	↗ HMD < Tablet	↘ HMD > Tablet			
55	HMD	Paper	↘ HMD > Paper	↗ HMD < Paper				