Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/GC03.20250143
Published: February 2, 2026

Usability, User Experience, and Workload of Augmented Reality vs. Paper Based instruction in Maintenance and Assembly Application

Faisal M. Alessa
Industrial Engineering Department
King Saud University
Riyadh, Saudi Arabia
fmalessa@ksu.edu.sa

Abstract

Augmented reality-based training is progressively utilized in the business to instruct personnel in a safe and efficient manner for maintenance activities. This study aims to examine the impact of HoloLens AR-based instructions versus paper-based instructions on performance, psychological effects, perceived workload, usability, and user experience in a maintenance task. Task completion time was used to evaluate user performance. The NASA-Task Load Index (NASA-TLX) questionnaire was utilized to evaluate workload. The usability of both instruction systems was evaluated using the extended System Usability Scale (SUS) questionnaire. The User Experience Questionnaire (UEQ) was employed to evaluate user satisfaction about both instructional methods. Twenty healthy males were enlisted for this investigation and assigned to two groups. The initial group conducted the experiment utilizing augmented reality-based education, whereas the subsequent group employed paper-based instructions. The results showed that performing maintenance task using AR instruction led to enhance user performance. Participants utilizing the AR were, on average, 17.5% more efficient than those following the printed instructions. Participants utilizing AR indicated an average 12% rise in perceived effort relative to the paper-based approach. The SUS indicates unfavorable evaluations for the paper-based education regarding both perceived utility and perceived ease of use. The AR was seen as highly beneficial and user-friendly.

Keywords
Augmented Reality (AR), maintenance task, perceived workload, user experience, and usability.

1. Introduction
Many industrial activities such as assembly and maintenance tasks, continue to be performed manually by humans. In order to carry out manual tasks, companies delivered written instructions for the operators. With the rising need for product variants, it is becoming progressively challenging to present paper-based work instructions in manual assembly and maintenance environments (Alhaag et al. 2024). Furthermore, locating the appropriate manual for a particular operation can be laborious and extends the process's lead time. Secondly, instructing new employees on these intricate assembly duties is frequently time-consuming and necessitates the participation of a senior colleague, which can be resource-intensive. In the current context, operators require more than basic text and image support, such as animations, films, or projections, for intricate assembly operations. Incorporating these skills onto conventional printed instructions is challenging. Industries and research organizations suggested virtual instructions as a substitute for conventional paper-based instructions. It is essential that systems designed to assist users with complicated tasks are assessed for their usability and utility to ensure effective, efficient, and rewarding user experiences.

Latest technologies such as Augmented Reality (AR) superimpose computer-generated images onto real environments. AR is characterized by the incorporation of virtual entities into the physical environment, allowing real-time human interaction with these entities (Alhaag et al. 2024). AR can be implemented in different ways, but the most common approach is utilizing augmented reality technology, such as AR glasses. AR creates value across industrial sectors, particularly in manufacturing, where it is applied in assembly work instructions, remote maintenance, quality assurance, and training (Alessa et al. 2023; Alhaag et al. 2024).

This study presents the design and assessment of an augmented reality maintenance guidance system, comparing it with conventional paper-based instructions. The AR maintenance instructions are implemented using a head-mounted display (HoloLens®). A comparison is conducted in terms of task completion time, overall workload, and system usability between paper-based and AR-based instructions. This approach enables a clearer understanding of the advantages and limitations of adopting new instructional systems to replace legacy methods. The study focuses on examining two instructional approaches for a maintenance training activity. The selected task involves the maintenance of a gearbox of a piston pump that incorporates both 3D-printed and conventional components. This requires a variety of activities, including fitting, screwing, and fastening—closely resembling a genuine shop floor environment. The two instructional systems under comparison are standard paper instructions and an augmented reality system delivered via a head-mounted display (HoloLens®).

1.1 Objectives

The aim of this study is to assess the efficiency and efficacy of augmented reality training compared to traditional paper instruction in maintenance applications. We assessed instructional efficiency by task completion time and instructional effectiveness via user usability, experience, and perceived workload using post-condition surveys. Our primary hypothesis posits that users would perceive augmented reality training as more effective, efficient, and pleasant compared to traditional paper-based instruction:

- H1: AR is more efficient than paper instruction.
- H2: AR is more effective than paper instruction.
- H3: AR is more satisfying than paper instruction.
- H4: AR is Less workload than paper instruction.

2. Literature Review

This section emphasizes research studies that have contrasted augmented reality with alternative modalities for task-based instructional applications. Specifically, it emphasizes how our efforts tackle a critical issue: the insufficient regard for the user. Various methodologies are employed in the industry to train workers, including paper-based instructions, e-learning, digital instructions, adaptive training, remote trainer assistance, virtual reality and augmented reality (Langendorf and Khalid 2025). Cognitive load and user experience are two essential concepts from the user's perspective. Cognitive load is a complex notion that refers to the demand imposed on a learner's cognitive system while doing a certain activity. The primary objective of cognitive demand in human–computer interaction is to improve the interface and interaction to reduce unnecessary cognitive burden, hence enabling more resources to be allocated to solving practical issues. User experience encompasses the perceptions and responses of individuals when engaging with or anticipating the usage of a system, product, or service, influencing their willingness to utilize a specific product or mode of engagement (Xie et al. 2025). Conducting user-centered study that incorporates cognitive demand and user experience to evaluate different interaction modes is crucial for their effective deployment. The System Usability Scale (SUS) is a commonly utilized standardized questionnaire for evaluating perceived usability. The SUS was employed to swiftly assess individuals' perceptions of the usability of the computer systems they were utilizing.

Recently, numerous researchers have concentrated on the deployment of immersive technologies to enhance real-time immersive experiences for student participants in industrial contexts (Eswaran and Bahubalendruni 2022). Numerous researches on immersive technologies have been published recently, detailing the technical hurdles, current advancements, future trends, intricate design, and utilized tools for applying immersive technology in industrial settings. The execution of assembly duties for the intricate product necessitates enhanced guidance and aid strategies in the real-time industrial environment (Makris et al. 2013; Lotsaris et al. 2021). This section provides a concise overview of significant AR/VR immersive technology. Porter et al., (2010) executed a substantial research investigation on the effects of immersive augmented reality experiences on hand-held displays (mobile), reporting a substantial impact on flexibility and ergonomic considerations. Hou et al., (2013) compared augmented reality with paper instructions in a Lego™ assembly activity. The findings indicated a reduced job completion time and diminished burden for the AR group.

Loch et al., (2016) examined augmented reality and video for assembly assistance utilizing Lego™. Each instructional style was assessed objectively for precision (task completion (TC)) and efficacy (time-to-completion (TTC)). The results demonstrated a reduction in time to completion (TTC) and cognitive effort, alongside a statistically significant drop in mistake frequency during Augmented Reality (AR) training. The authors demonstrated via the use of multiple regression analysis that 'perceived satisfaction' strongly affects 'use intention'. The minimal influence of 'perceived external control' was disregarded because of the experimental environment.

Vanneste et al., (2020) conducted an investigation of augmented reality possibilities for assembly duties, comparing it with conventional verbal and paper-based approaches. The authors examined performance concerning dependency, productivity, and completion time. Other researchers performed a comprehensive investigation on task information dissemination and the efficacy of HMD for displaying and analyzing mistake rates and task completion times (Werrlich, Daniel, et al., 2018; Werrlich, Lorber, et al., 2018; Werrlich, Nguyen, et al., 2018). Koumaditis et al., (2019) conducted an experimental investigation on augmented reality-assisted assembly tasks, recording differences in task execution time, productivity, and supplementary efforts. Polvi et al., (2018) conducted a comparable experiment that analyzed the differences in task completion times with and without augmented reality support. The current investigations employed AR video and animation-based instructional visualization for workers to adhere to task specifics. The significance of augmented reality mapping with tangible objects is overlooked by researchers in their experimental experiments. Therefore, the efficacy of the innovative AR mapping-based method, in conjunction with alternative instructional delivery modes for maintenance applications requires further investigation. The results of the suggested mapping method must be validated regarding productivity, execution time, and cognitive load.

Extensive investigations have been conducted on the development of AR applications and the evaluation of suitable visual cues for user interface (J. B. Alves et al. 2021; M. Alves et al. 2024; Egger and Masood 2020; Siriborvornratanakul 2018; Young and Rai 2021; Zhu, Ong, and Nee 2014; Belsito, Franciosi, and Di Pasquale 2024). AR applications have been developed for several manufacturing functions, including repair, maintenance, and assembly tasks. This study presents the developing and assessment of augmented reality-based maintenance instructions, comparing them with traditional paper-based instructions. This study aims to investigate two instructional methods for maintenance training activities.

3. Methods

3.1 Apparatus

The maintenance operation was executed on the gearbox of a piston pump system. This piston pump is a component of the “GUNT” Practice Line for assembly, maintenance, and repair, intended for technical training in universities and companies (set: MT 184). The instruments necessary for executing the maintenance operation on the piston pump included a screwdriver, wrench, soft hammer, jig, and bearing puller. Participants were guided in completing maintenance activity using two instructional methods: Microsoft HoloLens and paper instructions. A Microsoft HoloLens First Generation served as an augmented reality device to provide the maintenance guidelines. The paper instruction for the maintenance task was created using the manual supplied by the manufacturer (GUNT company). The gearbox of the piston pump was utilized as the case study due to its intricate structure and numerous components, making it appropriate for higher education students. Students can comprehend the functionalities and investigate the intricate geometry of each component within the pump. The created application fulfills three primary features of augmented reality (AR): The application improves students' experience by facilitating the exchange of contextual information. The digital model is presented upon receiving the actual educational material. Secondly, the digital content is amenable to real-time interaction. Digital material can be activated with visual attention and hand gesture commands. A command manual is created and presented at the bottom of the holographic screen for pupils to communicate with the system in an alternative manner. Third, the contents are spatially organized and associated with the three-dimensional space. Students may interact with the digital content in conjunction with the physical surroundings, such as positioning the model on a table and engaging with tangible things.

3.2 Participants

Twenty volunteers were recruited, with a mean age of 32.12 (SD=2.45). All participants were alumni of King Saud University in Saudi Arabia. To mitigate potential bias, the selected subjects possessed no prior familiarity with

Microsoft Hellenes. Participants possessed normal visual acuity (20/20) and exhibited no color blindness. The corrected visual acuity achieved with contact lenses was deemed satisfactory for the experiment. The study received permission from the Human Participants Review Sub-committee of the Institutional Review Board of King Khalid University Hospital, College of Medicine, King Saud University (ethical approval code: E-19-4467).

3.3 Experimental Design

The study aims to investigate the instruction effectiveness of AR instructions presented through a HoloLens for maintenance and assembly tasks. This study used a between-experimental design. The independent factors were instruction methods (paper vs. AR). Between-groups comparisons were conducted across instruction methods to compare measures between training delivered through paper and AR instructions. Dependent variables were time to completion, psychological evaluation scores, perceived workload (NASA Tlx), user experience (UEQ), and user usability (System Usability Scale).

3.4 Building Instruction

Figure 1 presents a comprehensive framework for this study, examining the two instruction methods for industrial maintenance tasks, the specifics of laboratory experiments, and the instructional equipment utilized. The entire work is split into three sections: a) Selection of various instructional methods, b) Creation of instructions for the chosen methods, and c) Implementation of laboratory experiments and assessment of performance. A methodical methodology for creating instructions is examined. Furthermore, the specifications of the hardware and software employed in the proposed visualization technique are addressed.

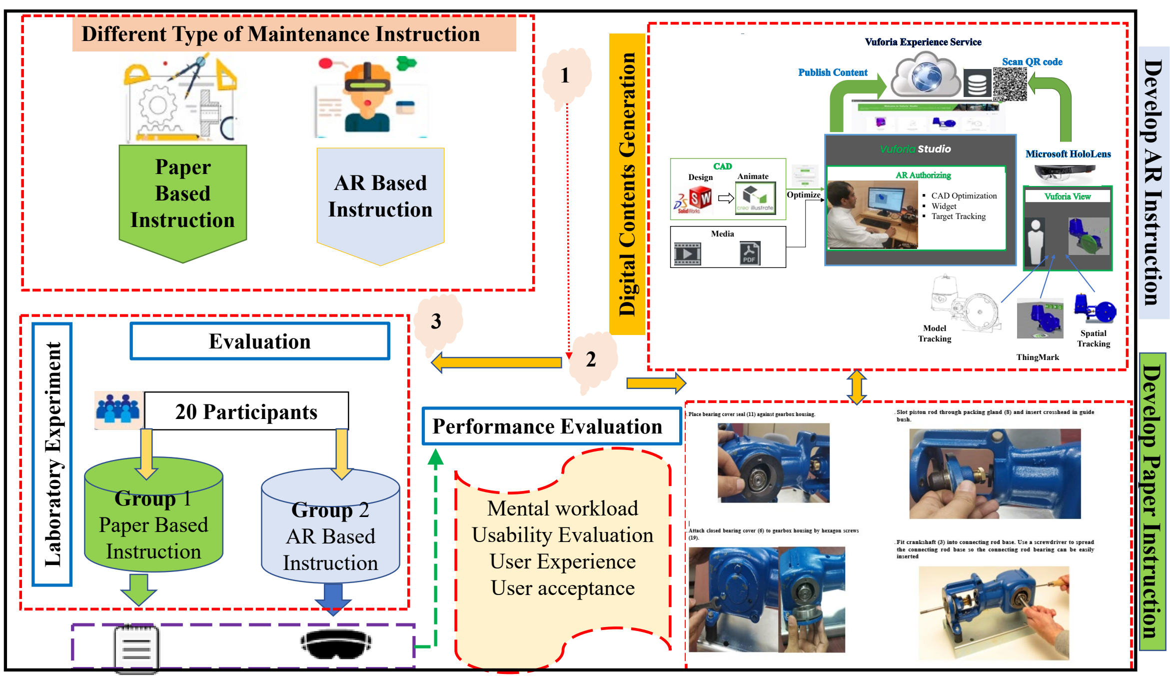
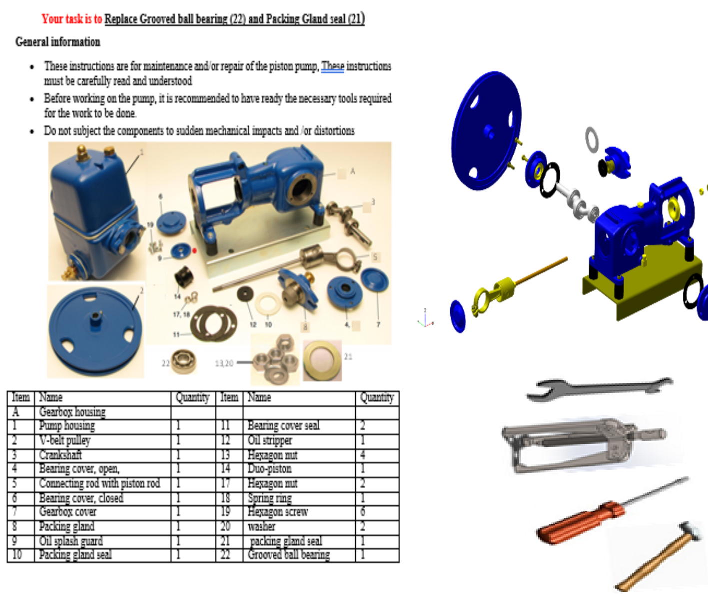


Figure. 1 Detailed framework of this study.

3.4.1 Building Paper Based Instructions

The manuals were derived from the piston pump manual utilized for troubleshooting and repair. The maintenance task of fixing a piston pump's gearbox was employed in this investigation, mirroring the tasks utilized in the research conducted by (Alhaag et al. 2022). In order to carry out their tasks, users may refer to the document that details the procedural maintenance activity. The bill of materials, various CAD representations of key components in two dimensions, three-dimensional visualizations of the piston pump, and necessary data about the assembly procedure are all included in the paper. The printed instructions were provided on A4 sheets, with one page allocated for every two steps. Each phase comprised images, textual information, and requisite equipment. The printed instructions comprised 22 pages and 40 steps. On the first page, there was a listing of all of the components that had been disassembled, along with the code numbers for each component and the objective of the operation, as illustrated in

Figure 2. Additionally, Figure 3 depicts the assembled component to improve workers' understanding of the assembly task performance. Pages 2 to 20 each included two of the 40 stages. The last page showed the whole installation of the gearbox, as illustrated in Figure 4. Conventional equipment, such as an open-ended wrench, screwdriver, puller, jig, and soft hammer, were used by the participants in order to do the job of fixing a piston pump gearbox and testing all of the seals.



Your task is to Replace Grooved ball bearing (22) and Packing Gland seal (21)

General information

- These instructions are for maintenance and/or repair of the piston pump, These instructions must be carefully read and understood
- Before working on the pump, it is recommended to have ready the necessary tools required for the work to be done.
- Do not subject the components to sudden mechanical impacts and /or distortions

Item	Name	Quantity	Item	Name	Quantity
A	Gearbox housing				
1	Pump housing	1	11	Bearing cover seal	2
2	V-belt pulley	1	12	Oil stripper	1
3	Crankshaft	1	13	Hexagon nut	4
4	Bearing cover, open,	1	14	Duo-piston	1
5	Connecting rod with piston rod	1	17	Hexagon nut	2
6	Bearing cover, closed	1	18	Spring ring	1
7	Gearbox cover	1	19	Hexagon screw	6
8	Packing gland	1	20	washer	2
9	Oil splash guard	1	21	packing gland seal	1
10	Packing gland seal	1	22	Grooved ball bearing	1

Figure. 2 Paper base instruction - Piston pump components and tool used.

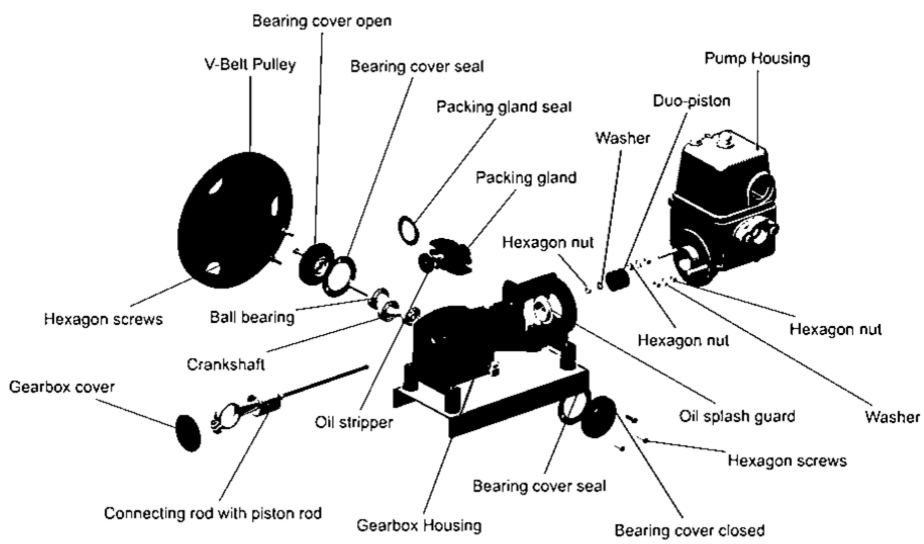


Figure. 3 Piston pump exploded view.

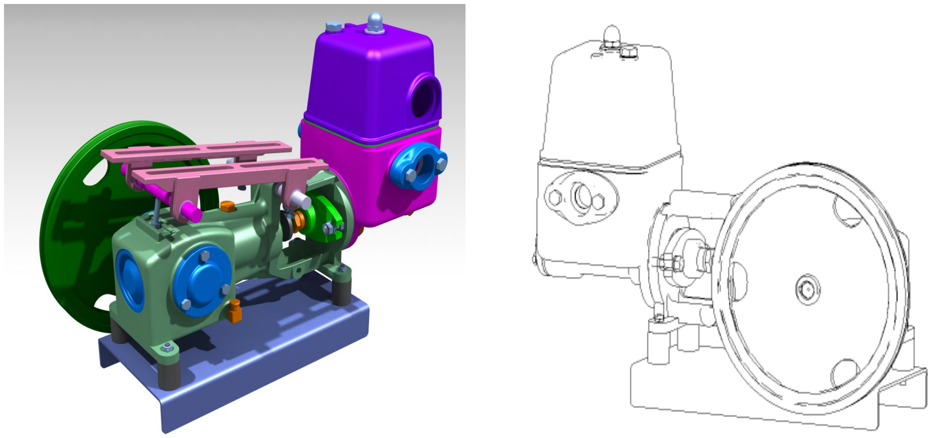


Figure. 4 piston pump fully assembled

The maintenance procedure comprised disconnecting the piston pump's gearbox and arranging all components on the table, followed through the process of removing the grooved ball bearing from the crankshaft and replacing it with a new one. The replacement of the grooved ball bearing is a difficult work that requires accuracy in addition to a variety of specialized hand tools. It also involves a number of subtasks that include several subtasks. To accomplish this, first place a puller over the grooved ball bearing. Then, evenly press the hook arms onto the bearing. To center the jig, adjust the thread against the shaft end. Finally, carefully turn the adjuster thread to detach the grooved ball bearing from the crankshaft. The process included using a jig and a pliable hammer to securely fasten the grooved ball bearing onto the crankshaft. The next step was to reassemble the gearbox of the piston pump, examine the packing gland seal, and install the new grooved ball bearing. Figure 5 shows written instructions for removing and replacing the crankshaft grooved ball bearing. Figure 6 illustrates that participants executed maintenance activities according to written instructions.

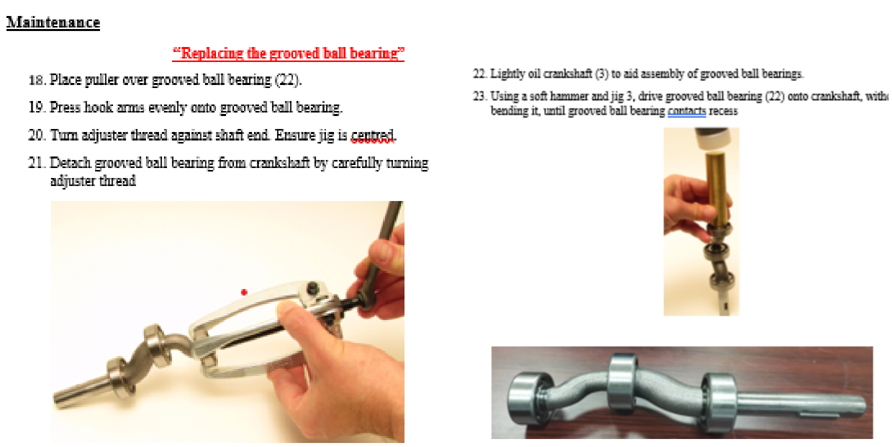


Figure. 5 One step of paper Based Instruction for replacing the grooved ball bearing

Figure. 6 Participants performed maintenance task based on paper instruction.

3.4.2 Building AR instruction

The augmented reality advice designed to assist participants in the routine maintenance of the piston pump has been developed and evaluated. Accompanying the user's natural vision with visual cues, directional arrows, and animated sequences, the augmented reality assistance utilizes a tracked head-mounted display to make job identification, localization, and execution easier. Figures 7 and 8 delineate the workflow for the development of the augmented reality application. The first step in creating the augmented reality instructions was to use SolidWorks software to build 1:1 CAD models of the piston pump. Afterwards, the augmented reality experience's sequence animations were created and tweaked using PTC Creo Illustrate. Building the 3D instructions for maintenance tasks included using visual cues and animations for installation, repair, and disassembly. Each step of the process has its own set of sequential instructions that make use of various motion effects, such as fly-in/out, fade-in/out, shaking, pulse, unscrew, and others. The effects are added to each step separately, and then all the stages of an instruction are combined into one smooth motion. Also included were supplemental materials including animated wrenches, pullers, arrows, and screwdrivers. Subsequently, the PLZ file containing the 3D guidelines was made available for use in creating the augmented reality experience using Vuforia Studio, as illustrated in Figure 8.

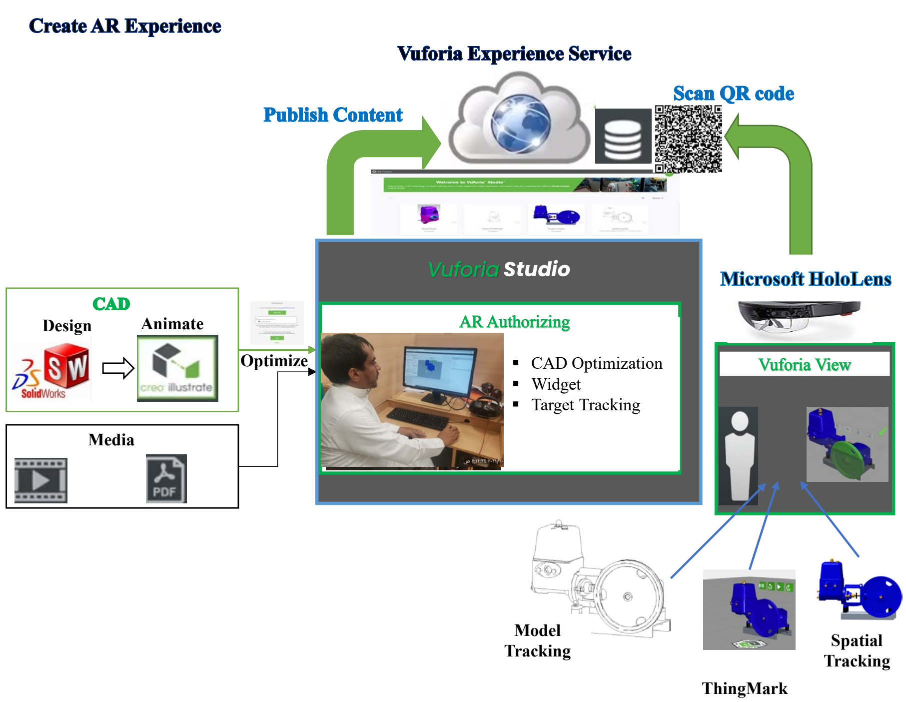


Figure. 7 Workflows for the development of the AR application

With Vuforia Studio, an AR software for the web, users can create and share augmented reality experiences, such as step-by-step work instructions and the ability to superimpose digital material over physical tools to provide more context.; refer to Figure 9. In addition, the apps' core features were developed using the computer language JavaScript. The experiences developed for maintenance activities were subsequently uploaded to the cloud-based "Vuforia Experience Service" via rapid response (QR) codes and ThingMarks. The Vuforia Experience Service is a server that stores and distributes published augmented reality experiences, enabling users to access them on Microsoft HoloLens 1 devices via the Vuforia View application.

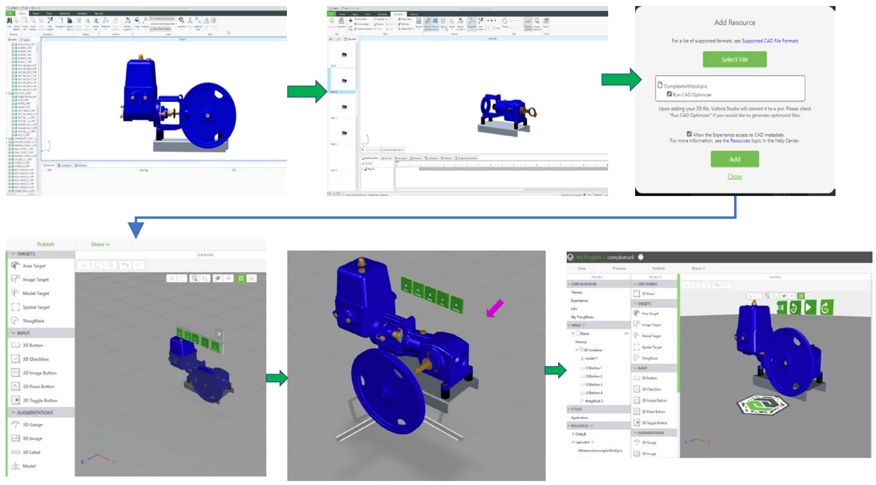


Figure 8. Creating a sequence by Creo Illustrate

The Vuforia View application on Microsoft HoloLens was utilized to display the augmented reality instructions by scanning a designated QR code, while a thing marker was positioned at a distance of 130 cm from the participant in the maintenance area. Upon identification of the object by Mark, Microsoft HoloLens superimposed a 3D CAD model of the designated task into the user's environment. This solution provided maintenance steps as 3D animated models within the augmented reality environment. Participants can engage with the 3D material and manipulate the animation playback by focusing their attention on the virtual buttons (e.g., play, next step, previous step, reset, task names) positioned above the 3D model and executing a pinching action with their fingers. Upon clicking the play button, a visual, step-by-step animated sequence for the chosen job is presented. Figure 10 illustrates an instance of augmented reality teaching for a single step (removing the dual-piston during gearbox maintenance). Figure 11 illustrates a participant engaging with augmented reality.

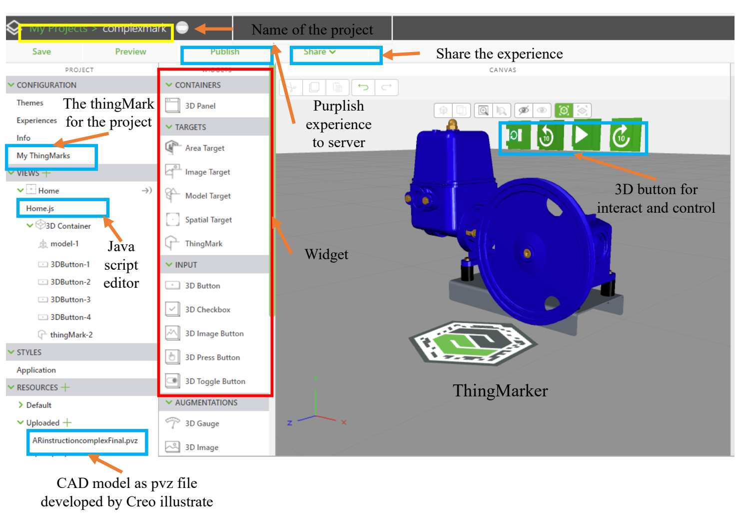


Figure. 9 Create AR by Vuforia Studio

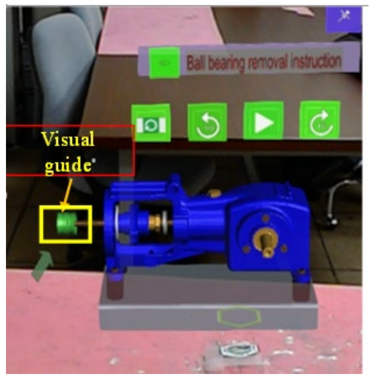


Figure.10 Digital content displayed by AR

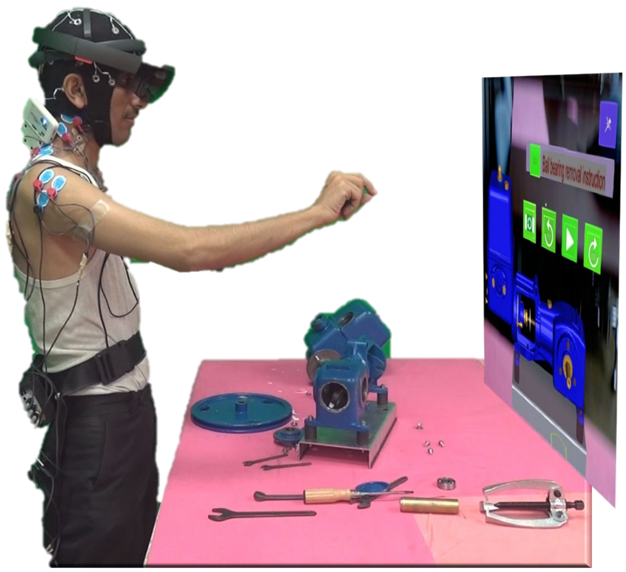


Figure. 11 One participant interact with AR

3.5 Experimental Setup and Procedure

The research was carried out in the maintenance laboratory of the Industrial Engineering Department. The laboratory's table center recorded a temperature of 21.4°C, humidity of 18.2%, and light intensity of 392.5 lux. The testing zone was protected from vibrations and strong scents during the task. At King Saud University, an invitation to participate was distributed and distributed to all who were interested. During the announcement, participants were given the opportunity to take part in a pre-test session, which was then followed by a test session that lasted between one and two hours. Upon arrival, participants were provided with a detailed explanation of the experiment's purpose, followed by the completion of the screening process. The health survey results were evaluated and approved; upon acceptance, Participants were asked to fill out a demographics questionnaire and sign a permission form. Participants were also asked to answer questions about their present experience with augmented reality and provide background information. Their anthropometric information was then collected.

The next step was to teach each group how to use their designated instructional method after randomly assigning them to one of two groups. The initial group was provided with AR-based instructions utilizing a Microsoft HoloLens 1 to present pertinent maintenance guidelines and assist the participant in executing the maintenance tasks. Participants in the AR-based instruction group were allotted 10 to 15 minutes to become acquainted with the HoloLens, including its handling, adjustment, and interaction. Participants were acquainted with the use of the toolbox, including tools such as a screwdriver, bearing puller, and jig. The second group received A4 maintenance guidance to help them complete the duties. The paper-based leading group learned toolbox use and document interpretation. Participants were told to leave the lab and return the next day to begin official training after demonstrating competency with the Microsoft HoloLens or paper instructions.

The next day, the participant went back to the lab to finish the repair jobs using the same method of instruction as the day before, whether it was utilizing augmented reality or paper instructions. Members of the first group, who received instructions via augmented reality, were given the task of doing the maintenance on this particular day. Group 2 participants used paper-based instructions and replicated the processes of group 1. Participants were invited to fill out surveys that assessed their subjective views of the task and approach. The surveys focused on psychological impacts, perceived workload, usability (SUS), and user experience (UQE). After the training activities, the time was logged.

4. Data Collection

4.1 Subjective Evaluation Questionnaires

The NASA-Task Load Index (NASA-TLX) questionnaire was utilized to assess workload (Hart and Staveland 1988). The NASA-TLX measures the six aspects of perceived workload: mental demand, physical demand, performance demand, effort, and frustration. It is a self-report assessment. A horizontal line with two endpoints—"low"(0) and "high" (100)—represented each of the six one-item scales. Table 1 shows the psychological data that the participants were asked to provide in response to a questionnaire that included fifteen items, all of which may be answered positively or negatively. The participants evaluated the instructional methods employed by assigning scores from −3 to +3 for each question. The usability of both instruction systems was evaluated using the extended System Usability Scale (SUS) questionnaire, as outlined by Bangor (Brooke 1996). All participants were presented with ten SUS questions followed by an additional eleventh question in sequence. Participants rated each statement on a five-point Likert scale, from strongly disagree to strongly agree. Ultimately, all participants were requested to evaluate the overall usability of the system. A rating of zero points indicates an awful score, one point reflects a poor score, two points signify an acceptable rating, three points denote a satisfactory rating, and four points represent an excellent rating.

Table 1 Psychological questionnaire

Items	Negative side 1-3	Score(0-7)	Positive side 4-7
Ease of Task Performance			
1. The maintenance task was	Difficult to achieve		Easy to achieve
2. The object was	complicated		Simple
3. The task procedure was	complicated		Simple
4. The position to be handled was	Complicated		simple
Stability:			
5. During the task	I was tensed		I was not tensed
6. After the task	I felt tired		I did not feel tired
Satisfaction			
7. The task was	Boring		enjoyable
8. Using the manual was	Uncomfortable		Comfortable
9. The task was	Easy to concentrate on		Not Easy to concentrate on
10. During the task	I didn't positively use the manual		I positively use the manual
11. After the task	I did not feel satisfied		I feel satisfied
Ease of manual use			
12. The manual was	Difficult to understand		Easy to understand
13. The object and manual were	Difficult to check		Easy to check
14. The structure of the manual was	Complicated		Simple
15. The detail of the manual was	Difficult to see		Easy to see

4.2 Data Analysis

We used SPSS Statistics (version 23.1; IBM Inc.) to carry out the statistical analysis. A significance threshold of 0.05 was determined, which is a Type I error. The statistical analysis results were confirmed to be reliable by validating the design assumptions, which included data continuity, homogeneity of variance, and normality. In order to validate the normalcy assumption, Kolmogorov-Smirnov tests were performed. To compare the levels of the various methods of instruction, we used an independent sample t-test. The amount of time it took to complete the task as well as the SUS rating were the dependent variables. Each question on the SUS had a possible score between 1 and 5. The whole score is computed. Not only was the overall score taken into consideration, but so were the individual scores for questions 3, 4, and 10, respectively, which dealt with the system's simplicity of use, the need of technical help, and the amount of knowledge required to utilize it. Every dependent variable had descriptive statistics computed, such as the mean and standard deviation.

5. Results and Discussion

5.1 Task Completion Time

A notable difference existed in the job completion time across conditions. Figure 12 shows the average completion times together with 95% confidence intervals (CI). Participants using the AR were, on average, 17.5% more efficient than those following the printed instructions. The results of the two-sampled t-test reveal a significant difference between the paper instruction (M= 1725.86 seconds, SD= 386.92) and the AR (M= 1468.89, SD= 93.32), ($P<0.05$).

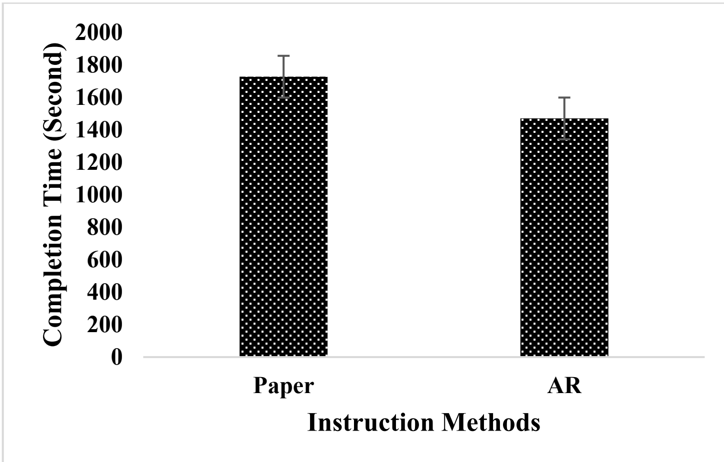


Figure. 12 task completion time across condition

5.2 Perceived Workload

We questioned the participants about their assessments of the physical and mental challenges, as well as the frustrations related to each lesson. The results indicated that individuals using augmented reality had, on average, a 12% increase in perceived effort relative to the paper-based approach. The results of the two-sample t-test demonstrate a significant difference between the paper teaching (M=63.833, SD=9.46) and the augmented reality (AR) method (M=71.48, SD=8.19, $P<0.05$), Figure 13.

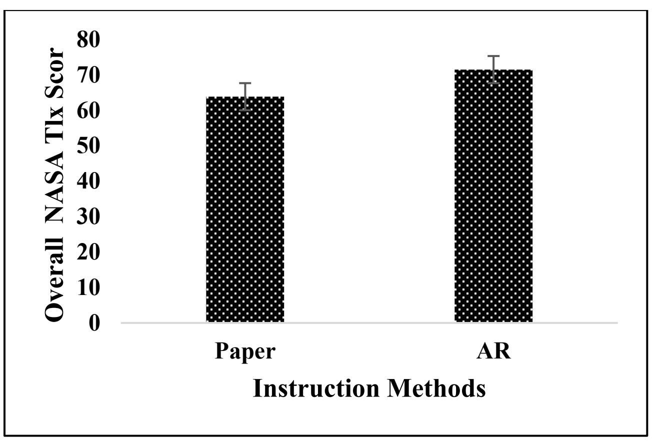


Figure. 13 Perceived workloads across condition

5.3 Psychological Effects

AR instruction was easy to understand, the structure of the instruction was simple in AR and complicate in paper and the detail of the instruction was easy to use in AR as compared to paper instruction. no difference between AR and paper-based instruction in Ease of Task Performance. Stability, satisfaction, and ease of instruction use were significantly different between the two instruction methods, $p<0.005$, Figure 14.

Scale	Augmented Reality Instruction						Paper Base Instruction						
	Mean	STD	N	Confidence	Confidence Interval		Mean	STD	N	Confidence	Confidence Interval		0.05
Ease of Task Performance	0.35	2.29	10	1.42	-1.07	1.77	0.43	1.39	10	0.86	-0.44	1.29	0.9014
Stability	1.70	1.31	10	0.81	0.89	2.51	0.15	1.16	10	0.72	-0.57	0.87	**0.0003**
Satisfaction	1.83	0.42	10	0.26	1.57	2.09	1.08	1.02	10	0.63	0.45	1.71	**0.0053**
Ease of Instruction use	2.75	0.33	10	0.21	2.54	2.96	1.56	1.16	10	0.72	0.85	2.28	**0.0002**

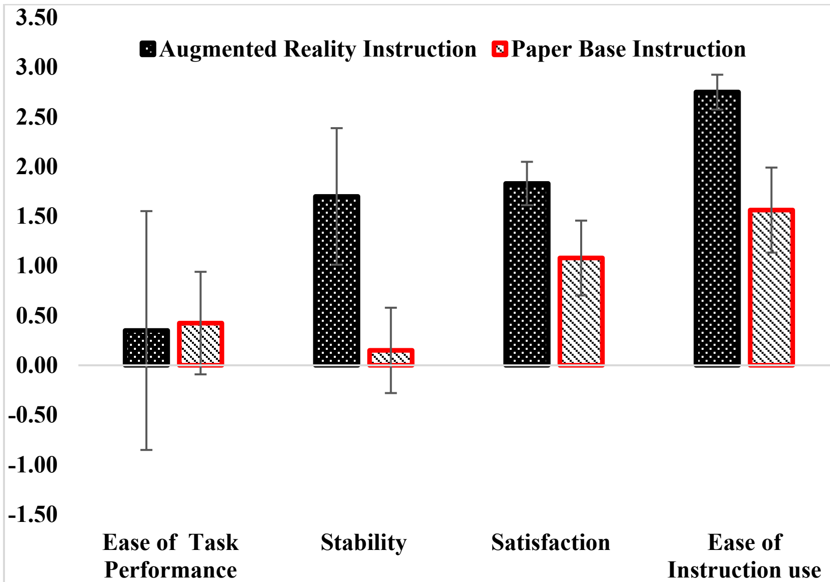


Figure. 14 Psychological Effects across condition

5.4 Perceived Usability

The SUS was administered to participants using five-point scales, with answer possibilities ranging from 1 (labeled "Strongly disagree") to 5 (labeled "Strongly agree"). If a participant does not react to an item, award it a score of 3 (the midpoint of the rating scale). Upon completion, ascertain the score contribution of each item, which will vary from 0 to 4. For favorably phrased questions (1, 3, 5, 7, and 9), the score contribution is calculated as the scale position minus 1. Deducting 5 from the item's scale position yields the negative phrase for items 2, 4, 6, 8, and 10. Multiply the sum of the item scores by 2.5 to get the total SUS score. That is to say, take the raw score and remove 1 for items with an odd number of digits, and 5 for items with an even number of digits. The standard SUS score is the sum of all adjusted scores, which is then multiplied by 2.5. The SUS scale goes from 0 to 100, with 2.5 point increments being the standard. Perceived usefulness and simplicity of use were both negatively impacted by paper-based teaching, according to the SUS. The augmented reality was seen as very beneficial and user-friendly. A high SUS score indicates excellent usability. The results demonstrate a substantial difference between the paper teaching (M=52.14, SD=15.61) and the augmented reality (AR) method (M=63.7, SD=11.19, P<0.05).

5.4 UEQ

The user experience questionnaire was used to conduct the evaluation of the user experience (Schrepp and Thomaschewski 2019; Schrepp, Hinderks, and Thomaschewski 2017). The UEQ, on the other hand, is comprised of a total of twenty-six questions and includes classifications of pragmatic quality such as perspicuity, efficiency, and dependability, as well as classifications of hedonic quality such as novelty and stimulation. The UEQ is not only used to evaluate hedonic and pragmatic quality, but it also uses seven-point semantic differential items to evaluate both types of quality. There is no overall score that is provided by the UEQ; rather, there is a score that is tied to six individual areas. The calculation of the score for each category involves taking the average of the several elements that are included inside it. The value of each item may vary from -3 to 3, with extreme values representing notions that are diametrically opposed to one another, such as attractiveness and unattractiveness. The 26 things are categorized into 6 scales:

- Attractiveness: This aspect of the product describes the overall impression that it gives. Are consumers satisfied with it or dissatisfied with it? Is it appealing, pleasurable, or interesting to look at?
- Perspicuity: A description of how simple it is to get acquainted with the product. Would it be simple to learn? Does the product have a straightforward and clear description?
- Efficiency refers to the ability of users to complete their jobs without exerting extra effort. Is the engagement quick and effective in its process? How soon does the product respond to the input of the user?
- Dependability: Does the user get the impression that they are in command of the interaction? Is he or she able to anticipated the behavior of the system? When working with the product, does the user maintain a sense of self-assurance?
- Stimulation: Does the product provide a sense of excitement and is it compelling to use? It is a pleasure to use, right?
- Novelty: Does the product exhibit creative and inventive qualities? Is it able to attract the attention of the users?

Participants indicated that AR are more attractiveness, perspicuity, efficiency, dependability, stimulation and novelty than paper-based instruction, Figure 15.

UEQ Scales (Mean and Variance)

Scale	AR instruction						Paper Instruction						
	Mean	STD	N	Confidence	Confidence Interval		Mean	STD	N	Confidence	Confidence Interval		
Attractiveness	2.56	0.34	10	0.21	2.35	2.77	1.02	0.90	10	0.56	0.46	1.57	0.0000
Perspicuity	2.51	0.36	10	0.22	2.29	2.73	0.69	1.15	10	0.71	-0.02	1.40	0.0000
Efficiency	2.05	0.43	10	0.27	1.78	2.32	1.24	0.80	10	0.50	0.74	1.73	0.0004
Dependability	2.50	0.69	10	0.43	2.07	2.93	0.75	0.65	10	0.41	0.34	1.16	0.0000
Stimulation	2.63	0.44	10	0.27	2.35	2.90	1.45	0.89	10	0.55	0.90	2.00	0.0000
Novelty	2.43	0.49	10	0.31	2.12	2.73	0.31	0.88	10	0.54	-0.23	0.86	0.0000

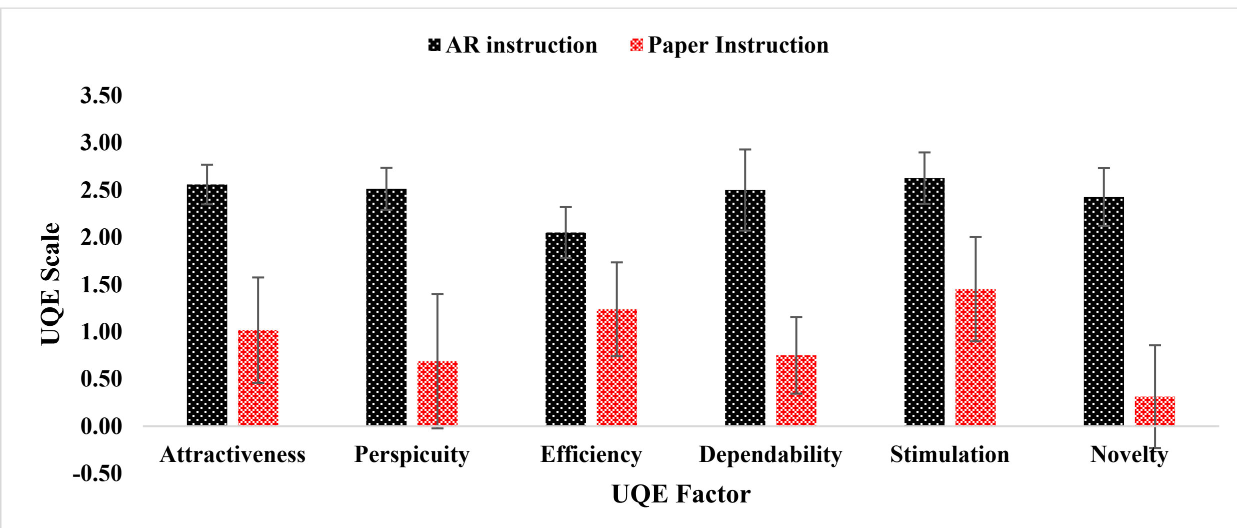


Figure. 15 User experience across condition

Scores that fall within the range of -0.8 to 0.8 are often indicative of a neutral rating, but values that exceed 0.8 are indicative of a positive evaluation, and values that fall below -0.8 are indicative of a negative review. The findings that were obtained are remarkable due to the fact that all of the categories got mean scores that were either greater than or very near to 2, and also due to the fact that severe UEQ values are very irregular. In addition, the lower limits of all confidence intervals for each category are considerably higher than the minimal level that was specified to be recognized as positive assessments ($p = 0.05$). UX results in terms of UEQ have been benchmarked using a dataset that contains data from 9905 responses that correspond to 246 studies (Hinderks, Schrepp, and Thomaschewski 2018; Schrepp, Hinderks, and Thomaschewski 2017). However, due to the fact that instruction categories have not been taken into consideration, this benchmarking can only be used as a first indicator to evaluate the user experience of the system that is being investigated. According to the results of the benchmarking, which are shown in Figure 16 and Figure 17, AR instruction ranks above average (that is, in the top 10%) in each of the six areas.

	AR instruction			Paper Instruction		
Scale	**Mean**	**Comparison to benchmark**	**Interpretation**	**Mean**	**Comparison to benchmark**	**Interpretation**
Attractiveness	2.56	Excellent	Upwards of 10% for optimal performance	1.02	Below average	50% of outcomes improved, 25% of outcomes deteriorated
Perspicuity	2.51	Excellent	Upwards of 10% for optimal performance	0.69	Bad	Within the lowest 25% of outcomes
Efficiency	2.05	Excellent	Upwards of 10% for optimal performance	1.24	Above Average	25% of outcomes improved, 50% of outcomes deteriorated
Dependability	2.50	Excellent	Upwards of 10% for optimal performance	0.75	Bad	Within the lowest 25% of outcomes
Stimulation	2.63	Excellent	Upwards of 10% for optimal performance	1.45	Good	10% of outcomes improved, 75% of outcomes deteriorated
Novelty	2.43	Excellent	Upwards of 10% for optimal performance	0.31	Below Average	50% of outcomes improved, 25% of outcomes deteriorated

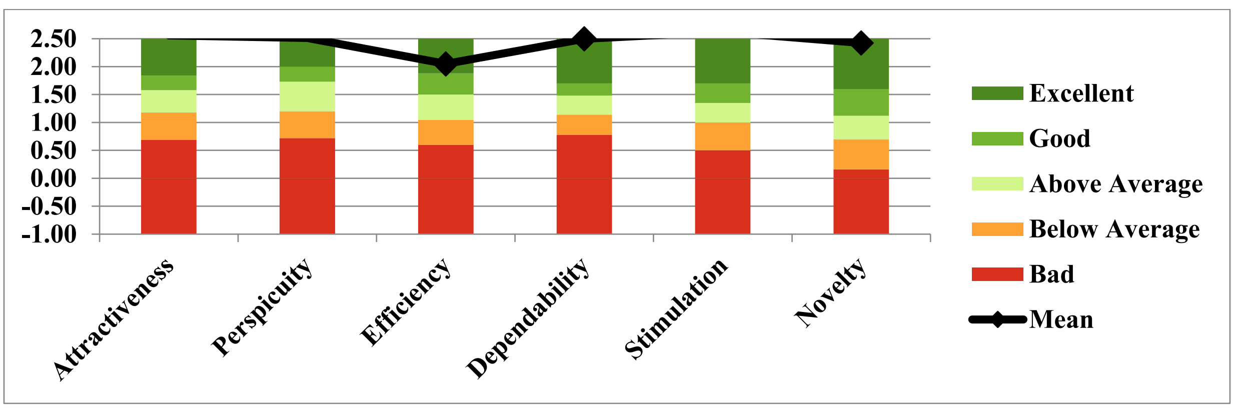


Figure. 16 User experience of AR based instruction compared to benchmark

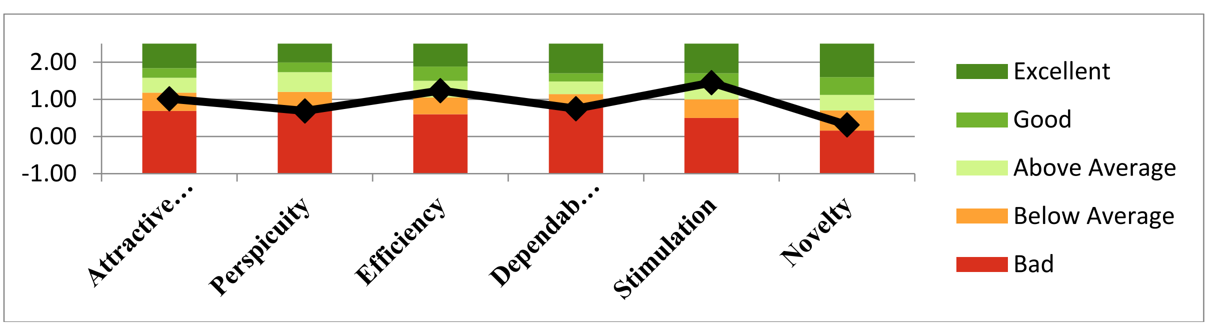


Figure. 17 User experience of paper-based instruction compared to benchmark

The AR demonstrates a practical and low-cost solution for presenting instruction. The presenting of instruction with AR was effective. The scales of the UEQ can be grouped into pragmatic quality (Perspicuity, Efficiency, Dependability) and hedonic quality (Stimulation, Originality). In contrast to hedonic quality, which focuses on factors unrelated to tasks, pragmatic quality identifies those that are directly relevant to completing the work at hand. You may find the average of the three hedonic and pragmatic quality metrics down below, as shown in Figure 18 and Figure 19.

Pragmatic and Hedonic Quality		
	Paper	AR
Attractiveness	1.02	2.56
Pragmatic Quality	0.89	2.35
Hedonic Quality	0.88	2.53

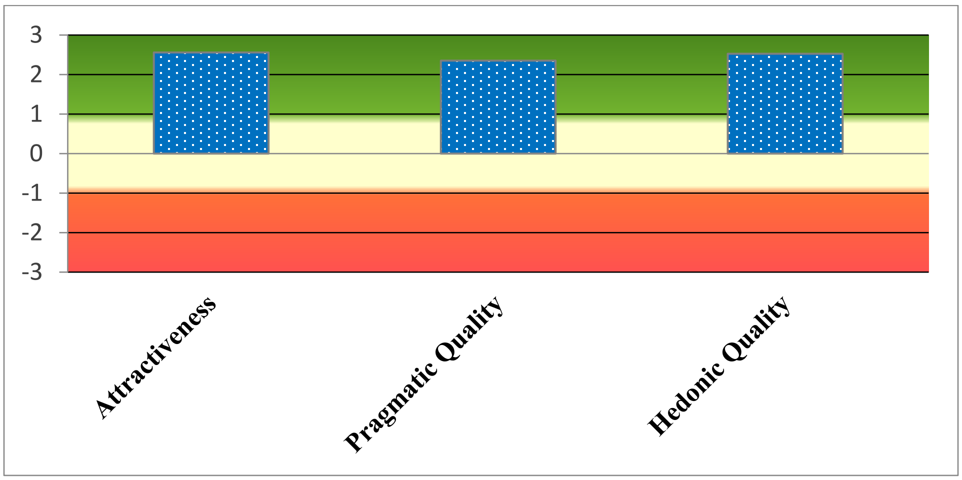


Figure. 18 UEQ grouped for AR based instruction

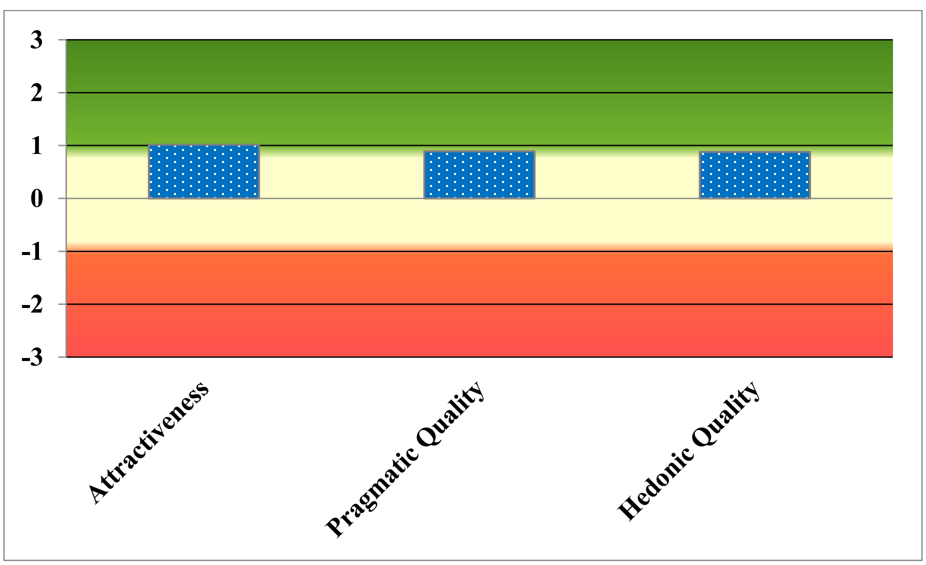


Figure. 19 UEQ grouped for paper Based instruction

UEQ does not provide a total UX score, but rather separate ratings for each aspect. KPI extension of the UEQ was used to evaluate overall score for the UX, Figure 20. This extended UEQ contains 6 additional questions, which ask for the importance of the UEQ scales for the participant. With this information a KPI can be calculated, i.e. a single number that can be interpreted as overall UX impression. Overall UX impression, For AR, KPI= 2.44 , SD=0.25, KPI= 0.91, SD=0.20.

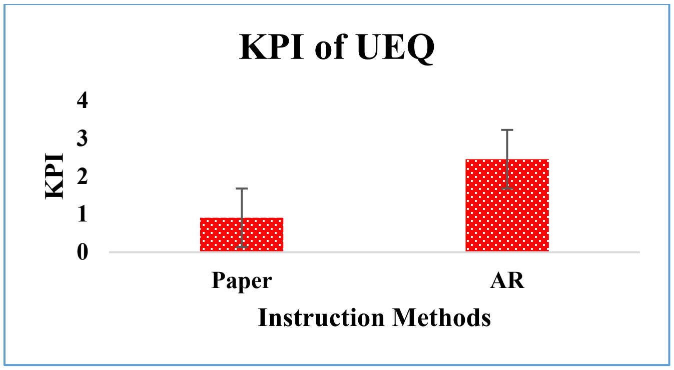


Figure. 20 overall score for the UX of AR Based Instruction paper instruction

Three main requirements of augmented reality are met by the created app: To start, by facilitating the sharing of contextual information, the application improves the experience for participants. After the marking is recognized, the digital model is shown. Second, people may engage with animated models in real-time thanks to the digital content. Gazing and hand motion instructions may be used to communicate digital material. In order to provide the players another means of signaling the system, virtual buttons were created and placed on top of the holographic screen. Third, the contents are connected with the 3D space and spatially mapped. As an example, users may put the virtual model on a table and see how it reacts when it hits actual things in the room. The main elements of HoloLens that allow users to interact with digital models are gazing (head movement tracking) and hand gestures.

5. Conclusion

This study presents the design and assessment of an augmented reality maintenance guidance system, comparing it with conventional paper-based instructions. The AR maintenance instructions are implemented using a head-mounted display (HoloLens®). A comparison is conducted in terms of task completion time, psychological effects, perceived workload, usability, and user experience between paper-based and AR-based instructions. This approach enables a clearer understanding of the advantages and limitations of adopting new instructional systems to replace legacy methods. Our primary hypothesis posits that users would perceive augmented reality training as more effective, efficient, and pleasant compared to traditional paper-based instruction. Twenty healthy males were enlisted for this investigation and assigned to two groups. The initial group conducted the experiment utilizing augmented reality-based education, whereas the subsequent group employed paper-based instructions. The results showed that performing maintenance task using AR instruction led to enhance user performance. Participants utilizing the AR were, on average, 17.5% more efficient than those following the printed instructions. Participants utilizing AR indicated an average 12% rise in perceived effort relative to the paper-based approach. The SUS indicates unfavorable evaluations for the paper-based education regarding both perceived utility and perceived ease of use. The AR was seen as highly beneficial and user-friendly.

Reference

Alessa, Faisal M, Mohammed H Alhaag, Ibrahim M Al-Harkan, Mohamed Z Ramadan, and Fahad M Alqahtani. "A Neurophysiological Evaluation of Cognitive Load during Augmented Reality Interactions in Various Industrial Maintenance and Assembly Tasks." *Sensors* 23 (18): 7698. 2023.

Alhaag, Mohammed H, Faisal M Alessa, Ibrahim M Al-harkan, Mustafa M Nasr, Mohamed Z Ramadan, and Saleem S AlSaleem. "Effects of Industrial Maintenance Task Complexity on Neck and Shoulder Muscle Activity During Augmented Reality Interactions." *Electronics* 13 (23): 4637. 2024.

Alhaag, Mohammed H, Mohamed Z Ramadan, Ibrahim M Al-harkan, Faisal M Alessa, Hisham Alkhalefah, Mustufa Haider Abidi, and Abdelaty E Sayed. "Determining the Fatigue Associated with Different Task Complexity during Maintenance Operations in Males Using Electromyography Features." *International Journal of Industrial Ergonomics* 88:103273. 2022.

Alves, Joao Bernardo, Bernardo Marques, Paulo Dias, and Beatriz Sousa Santos. "Using Augmented Reality for

Industrial Quality Assurance: A Shop Floor User Study." *The International Journal of Advanced Manufacturing Technology* 115 (1): 105–16. 2021.
Alves, Manuel, Tiago Ribeiro, Leonel Santos, Anabela Marto, Alexandrino Gonçalves, Carlos Rabadão, and Rogério Luís de C. Costa. "Evaluating Service-Based Privacy-Protection for Augmented Reality Applications." In *International Conference on Management of Digital*, 254–63. 2024.Springer.
Belsito, Chiara, Chiara Franciosi, and Valentina Di Pasquale. "The Impact of Augmented Reality Devices on Operator Performance in Manufacturing Contexts." *Management and Production Engineering Review*. 2024.
Brooke, John. "SUS-A Quick and Dirty Usability Scale." *Usability Evaluation in Industry* 189 (194): 4–7. 1996.
Egger, Johannes, and Tariq Masood. "Augmented Reality in Support of Intelligent Manufacturing--a Systematic Literature Review." *Computers & Industrial Engineering* 140:106195. 2020.
Eswaran, M, and M V A Raju Bahubalendruni. "Challenges and Opportunities on AR/VR Technologies for Manufacturing Systems in the Context of Industry 4.0: A State of the Art Review." *Journal of Manufacturing Systems* 65:260–78. 2022.
Hart, Sandra G, and Lowell E Staveland. "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research." *Advances in Psychology* 52:139–83. 1988.
Hinderks, Andreas, Martin Schrepp, and Jörg Thomaschewski. "A Benchmark for the Short Version of the User Experience Questionnaire." In *Proceedings of the 14th International Conference on Web Information Systems and Technologies-APMDWE*, 373–77. 2018. SciTePress.
Hou, Lei, Xiangyu Wang, Leonhard Bernold, and Peter E D Love.. "Using Animated Augmented Reality to Cognitively Guide Assembly." *Journal of Computing in Civil Engineering* 27 (5): 439–51. 2013
Koumaditis, Konstantinos, Sarune Venckute, Frederik S Jensen, and Francesco Chinello. "Immersive Training: Outcomes from Small Scale AR/VR Pilot-Studies." In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 1–5. 2019. IEEE.
Langendorf, Lasse Nielsen, and Md Saifuddin Khalid. "Systematic Literature Review on Usability and Training Outcomes of Using Digital Training Technologies in Industry." *Computers in Human Behavior Reports* 17:100604. 2025.
Loch, Frieder, Fabian Quint, and Iuliia Brishtel. "Comparing Video and Augmented Reality Assistance in Manual Assembly." In *2016 12th International Conference on Intelligent Environments (IE)*, 147–50. 2016.
Lotsaris, Konstantinos, Nikos Fousekis, Spyridon Koukas, Sotiris Aivaliotis, Niki Kousi, George Michalos, and Sotiris Makris. "Augmented Reality (AR) Based Framework for Supporting Human Workers in Flexible Manufacturing." *Procedia Cirp* 96:301–6. 2021.
Makris, Sotiris, George Pintzos, Loukas Rentzos, and George Chryssolouris. "Assembly Support Using AR Technology Based on Automatic Sequence Generation." *CIRP Annals* 62 (1): 9–12. 2013.
Polvi, Jarkko, Takafumi Taketomi, Atsunori Moteki, Toshiyuki Yoshitake, Toshiyuki Fukuoka, Goshiro Yamamoto, Christian Sandor, and Hirokazu Kato. "Handheld Guides in Inspection Tasks: Augmented Reality versus Picture." *IEEE Transactions on Visualization and Computer Graphics* 24 (7): 2118–28. 2018.
Porter, Shane R, Michael R Marner, Ross T Smith, Joanne E Zucco, and Bruce H Thomas. "Validating Spatial Augmented Reality for Interactive Rapid Prototyping." In *2010 IEEE International Symposium on Mixed and Augmented Reality*, 265–66. 2010. IEEE.
Schrepp, Martin, Andreas Hinderks, and Jorg Thomaschewski. "Construction of a Benchmark for the User Experience Questionnaire (UEQ)." *International Journal of Interactive Multimedia and Artificial Intelligence* 4 (4): 40–44. 2017.
Schrepp, Martin, and Jorg Thomaschewski. "Design and Validation of a Framework for the Creation of User Experience Questionnaires." *IJIMAI* 5 (7): 88–95. 2019.
Siriborvornratanakul, Thitirat. "Enhancing User Experiences of Mobile-based Augmented Reality via Spatial Augmented Reality: Designs and Architectures of Projector-camera Devices." *Advances in Multimedia* 2018 (1): 8194726. 2018.
Vanneste, Pieter, Yi Huang, Jung Yeon Park, Frederik Cornillie, Bart Decloedt, and Wim den Noortgate. "Cognitive Support for Assembly Operations by Means of Augmented Reality: An Exploratory Study." *International Journal of Human-Computer Studies* 143:102480. 2020.
Werrlich, Stefan, Austino Daniel, Alexandra Ginger, Phuc-Anh Nguyen, and Gunther Notni. "Comparing HMD-Based and Paper-Based Training." In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 134–42. 2018.
Werrlich, Stefan, Carolin Lorber, Phuc-Anh Nguyen, Carlos Emilio Franco Yanez, and Gunther Notni. "Assembly Training: Comparing the Effects of Head-Mounted Displays and Face-to-Face Training." In *International Conference on Virtual, Augmented and Mixed Reality*, 462–76. 2018.

Werrlich, Stefan, Phuc-Anh Nguyen, and Gunther Notni. “Evaluating the Training Transfer of Head-Mounted Display Based Training for Assembly Tasks.” In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*, 297–302. 2018.

Xie, Xiaojiao, Yao Wang, Yan Cui, Suihuai Yu, Dengkai Chen, and Jianjie Chu. “Evaluation of Cognitive Load and User Experience in Alternative Interaction Modes under Different Noise Degrees.” *Advanced Engineering Informatics* 65:103328. 2025.

Young, Crystal, and Rahul Rai. Automated Procedure Reconfiguration Framework for Augmented Reality-Guided Maintenance Applications.” *Journal of Computing and Information Science in Engineering* 21 (6): 61011. 2021. “

Zhu, J, S K Ong, and A Y C Nee. “A Context-Aware Augmented Reality System to Assist the Maintenance Operators.” *International Journal on Interactive Design and Manufacturing (IJIDeM)* 8 (4): 293–304. 2014.