An Overview of Evaluations Using Augmented Reality for Assembly Training Tasks

S. Werrlich, E. Eichstetter, K. Nitsche, G. Notni

Abstract—Augmented Reality (AR) is a strong growing research topic in different training domains such as medicine, sports, military, education and industrial use cases like assembly and maintenance tasks. AR claims to improve the efficiency and skill-transfer of training tasks. This paper gives a comprehensive overview of evaluations using AR for assembly and maintenance training tasks published between 1992 and 2017. We search in a structured way in four different online databases and get 862 results. We select 17 relevant articles focusing on evaluating AR-based training applications for assembly and maintenance tasks. This paper also indicates design guidelines which are necessary for creating a successful application for an AR-based training. We also present five scientific limitations in the field of AR-based training for assembly tasks. Finally, we show our approach to solve current research problems using Design Science Research (DSR).

Keywords—Assembly, augmented reality, survey, training.

I. INTRODUCTION

ASSEMBLY training of novice employees plays a crucial role in the car manufacturing industry. Processes and products become more and more complex, which means a higher responsibility and challenge for every single assembly employee. Therefore, a good assembly training can prevent errors and should be efficiently designed. A common method is face-to-face training where an expert is teaching the novice employee new assembly procedures. This method is time consuming and expensive because it requires an always available expert for the trainee. A new technology for training, which is discussed in several studies, is AR. AR claims to improve the efficiency and skill-transfer of assembly tasks by visualizing computer generated 3D information in the real environment (Fig. 1). There has been no structured literature review in that field so far and therefore, this paper aims to give a comprehensive overview of AR based assembly studies.

This paper focuses on AR-based training applications for assembly tasks, especially on evaluations and guidelines because we want to answer the following question: How to design an AR-based application for assembly training and which advantages result in the comparison to an expert based training?

We make the following contributions:

- In Section II, we describe our literature research process. Four online databases are used for the research using the keywords “augmented reality” and “training”. The search returned 862 papers, with a select 54 papers in the field of AR-based training for assembly and maintenance tasks.

Fig. 1 AR-based engine assembly training using a HMD

- In Section III, we present the research results and describe the papers found. We give background information about the participants in the studies and the results of the evaluations. Section III ends with the presentation of design guidelines for AR, especially for AR-based assembly and maintenance training tasks.

- Based on the existing research, we indicate five limitations in Section IV: the participants, the assembly tasks, the comparisons, the KPIs and the technology.

- In Section V, we describe our approach to solve the indicated research problems. The DSR framework is best suitable for our future research. The methodology to address the seven DSR guidelines is presented at the end of this section.

- The summary of the paper is given in Section VI.

II. LITERATURE RESEARCH PROCESS

Most current scientific research follows an unstructured approach; it is not systematic, and therefore, lacks in completeness. This paper presents research with a more structured approach. Before starting the search, we define our keywords, which are necessary for the research in online databases. We use the keywords “augmented reality” and “training” for our research and limit our research to abstracts,

S. Werrlich is a PhD candidate at BMW Group Munich, 80788 München. He is also with the Group for quality assurance and industrial image processing at the Technical University Ilmenau, 98693 Ilmenau, Germany (e-mail: Stefan.Werrlich@bmw.de).

E. Eichstetter is an intern at the BMW Group Munich, 80788 München. She is also with the Hochschule München, 80335 München, Germany (e-mail: Eichstetter@hm.edu).

K. Nitsche is the head of the quality management department for electrified powertrains at BMW Group Munich, 80788 Munich, Germany (e-mail: Kai.Nitsche@bmw.de).

G. Notni is professor and the head of the Group for quality assurance and industrial image processing, Technical University Ilmenau, 98693 Ilmenau, Germany (e-mail: Gunther.Notni@tu-ilmenau.de).
titles and keywords. We use online databases which are recommended in previous summaries for AR [1]. Those databases are: IEEE Xplore, ACM Digital Library, Science Direct and Web of Science. In IEEE Xplore and ACM Digital Library it is only possible to search either for abstracts, titles or keywords. A combined search is not possible. A combined search is possible in Science Direct and Web of Science. The search returned 862 papers (Fig. 2).

Fig. 2 Results for "augmented reality" AND training

It the next step, we read all document titles and abstracts and decided if the paper is relevant to the present research. We are interested in user studies and design guidelines for AR-based training because this is still an unexplored field [2]. This paper aims to create a status quo in order to present further research directions.

Scientific contributions from other domains like medicine [3], sports [4] and military [5] are excluded because this is not scope of this paper. Also excluded are paper from the field of virtual reality (VR) training [6]. Research contributions focusing only on the description of AR-based systems [7] (e.g. new tracking algorithms) for training but not on evaluation were excluded as well.

In [8], it is mentioned that a training application must differ from a guidance application because a step-by-step solution is not sufficient for training. An application with a strong guidance has the danger that the trainee becomes dependent on the support and therefore might be not able to perform the task without any assistant. Therefore, this paper is not focusing on AR-based guidance applications [9] for assembly or maintenance tasks.

We indicate 17 relevant papers, focusing on user evaluation and seven design recommendations for user interface design, including those for AR-based training applications. In order to create a comprehensive overview, it is necessary to check the references of the indicated paper. Based on the titles in the references, we decide if a paper fits in our category of AR based assembly and maintenance training. The indicated titles are searched in the online databases and added to our final database if an article fits the determined category. We present our results in the Section III.

III. AR-BASED ASSEMBLY TRAINING

In this section, we present user evaluations using AR for assembly and maintenance training tasks as well as design guidelines for AR-based training.

A. Evaluations

Rios et al. [10] compared a tablet-based solution with written instructions. Seventeen participants (all students) maintained a Boeing 737 engine bleed air system with the support of either a tablet or written instructions. Results showed a time improvement of 17% percent and 24% quality increase when using AR.

Another study by [11] focused on the benefits of AR in gender. A LEGO model was used for the assembly procedure. Twenty eight students participate in the study. They were divided into two groups with seven males and seven females in each group. The first group used a computer monitor to train the assembly process and the second group used a 3D printed manual. After finishing one LEGO assembly cycle, the participants had to repeat the assembly procedure without any assistance. The results show the AR helps both male and female trainees to learn the assembly procedure faster. AR-based training is also more effective for male and female trainees than a 3D manual.

Horejsi et al. [12] compared paper based instructions with AR, using a monitor. Twenty volunteers were asked to assemble a gully trap using paper based instructions. Another 20 participants assembled with AR support. Results indicate that trainees learn 43% faster, using AR.

A study from the late 90s [13] compared VR with AR as a training tool. In five groups with five students each, a water pump should be assembled. The study compares 2D drawing, a Desktop solution, stereoscopic glasses, immersive VR using a head-mounted display (HMD) and context free AR. Participants were faster when using AR compared to all other methods.

Valimont et al. [14] used four learning paradigms. The study compared a printed manual, a video tape, text annotations by mouse interaction and AR using a monitor. Participants were encouraged to learn the assembly procedure from an oil pump as well as the functions of single parts and the different part locations. The experiment indicated that AR performed better in the short-term and the long-term memory test compared to the other three methods.

Another study by Pathomaree et al. [15] focused on the skill transfer in assembly tasks, using a 2D and a 3D puzzle. Twenty participants were exposed to one of four experiments. Each participant was asked to assemble the 2D or 3D puzzle either with or without AR. All instructions were shown on a monitor. The assembly procedure was shown once. Trainees had to assemble the puzzle afterwards without any support. It was shown that AR helps to increase training efficiency. Trainees also needed less assembly steps to finish the task, which indicates a better skill-transfer when using AR.

Macchiarella et al. [16] measured the impact of short and long-term memory and compared different presentation modes: print-based, video-based, AR-based and interactive AR-based. A total of 96 participants, all students, were divided in four groups and had to study an aircraft oil pump with support from one presentation mode. Trainees were asked afterwards if they remember the names of major components as well as several functionalities in order to determine how much knowledge they acquired in the training session. The long-term recall test occurred seven days later where the same questions were asked. Results from the immediate and long-
term recall indicate that trainees forget less over time when using AR and interactive AR, compared to print and video-based learning.

Webel et al. [8] did a skill transfer evaluation and compared two learning approaches. Twenty participants, all technicians with at least two years of assembly experience attended to the study. They had to perform an assembly task with 25 steps. In Group 1, 10 trainees had to watch an instructional video showing the assembly task of an electro-mechanical actuator. The second group, consisting of 10 participants, performed the physical task with AR support using a tablet. All trainees did the training in the morning. They had to perform the same physical task without any assistance in the afternoon. The evaluation showed that trainees perform 20% slower and made 77% less un-solved errors when trained with AR.

Gavish et al. [17] compared VR-based assembly training with AR-based assembly training. 40 technicians were randomly assigned to four groups: VR training, AR training and two control groups. They used an electronic actuator with 25 assembly steps for the experiment. The VR and AR control group performed the assembly task twice with support of the VR training platform. The AR and control AR group performed the assembly task with the AR tablet-based platform just once. After training, participants had to perform the task again without any support, after a break of several hours. They determined that the VR and AR groups had significantly longer training times compared to the control groups. Results also indicated a higher user satisfaction and usability when using the AR platform for training. Furthermore, trainees made less assembly errors when using the AR platform.

Another study compared a video-see through solution with face-to-face training [18]. In this evaluation, an Oculus Rift VR Headset with a mounted pair of cameras to the HMD is used. Some 24 volunteers, twelve in each group, had to learn a maintenance task of an aircraft door. Group 1 had to observe what the trainer was doing and try to remember as much as possible. Group 2 used the HMD for the training. After training, trainees were asked to complete a knowledge retention test, which was a written multiple choice test with eight questions and a knowledge interpretation test. This test evaluated whether the assembly procedure was captured. Results indicated no significant difference for the knowledge retention test and no significant differences for the knowledge interpretation test between the two conditions. Trainees spent more time to complete the training when using the HMD.

Hou et al. [19] compared an animated AR-system, using a big screen, with a paper-based manual system. He used a Lego model for the experiments. 50 participants, all students, were conducted to compare both systems. None of them had experience in using AR. The first experiment aimed to compare the cognitive workload of both systems using the NASA-TLX questionnaire. They also measured the time to complete the assembly task and the number of errors (selection error, assembly sequential error, installation error and help from the AR system or manual). The first experiment indicated a shorter completion time using the AR system. Participants made also fewer errors with support of the AR system. A higher mental workload was measured when participants used the manual in comparison to the AR system for the assembly process. The second experiment aimed to measure the learning curve for the two scenarios in order to find a difference between the two systems. Participants had to assemble the Lego model once. They were encouraged to remember the assembly procedure. After 5 min relaxing, they had to assemble the model again without any assistance. The results show that the AR-based training reduces the learning curve and the number of errors.

In [20], a monitor-based system was used. For the assembly training task, they used 12 parts of a RV-10 aircraft. Seven engineering students (all male) with no prior knowledge of assembling took part in the experiment. All of them had to perform six assembly trials, measuring the time, errors and the questions users ask to the supervisor who was available at all times. Results indicate a shorter assembly time for an AR-assisted assembly compared to a traditional manual.

Peniche et al. [21] combined AR and VR. In the first training session, participants had to perform the assembly task in VR five times and continued in the second session with AR three times. The advantage of VR-based training is the save of resources because trainees learn with virtual parts. Additional, users continue the training with an AR-based system where they come into direct contact with the real parts. The task was to assemble a milling machine. Before they start the two training conditions, participants had to perform the task twice in order determine their initial skills. The control group performed the assembly task eight times using the conventional methodology. They measured the time to complete the task and found out that their AR system lowered the performance in session seven, when the switched from VR-based training to the AR-based training. It was also shown that VR was able to replace a major part of the training and transferred skills in a significant manner.

Rios et al. [22] compared three different methodologies for the transfer of knowledge of complex training task in aeronautical processes. He compared printed instructions with an audiovisual tool and an AR application using a laptop. The training duration was four hours. Equal to [20], they measured the time, errors and questions asked during the training to the supervisor. The same assembly task, as well as seven male engineering students, was used. Trainees needed more time completing the task when using the AR application. There were a 4.5min time difference between AR and the audiovisual tool. The longest training was with the traditional training method. The error and question comparison showed a difference between AR and the other methodologies. Trainees made up to four times fewer errors using the AR-system compared to the other methodologies and needed less help from the supervisor to complete the task.

Westefield et al. [23] used a HMD to compare a traditional AR-based system using a step-by-step guidance with an intelligent AR solution which gives feedback to the users. The aim was to determine the difference in knowledge retention between the two approaches. Sixteen randomly assigned
participants (11 male and five female students) performed one of the conditions. They had no experience with computer hardware assembly. All of them had to assemble five motherboard components. After performing the training once, participants had to repeat the task without any assistance. Only the names of the components were given and users had to correctly identify and install them. The completion time and errors were also measured during the process. There were no significant differences in errors or completion time comparing both systems. Trainees who learned with the intelligent AR system scored 25% higher in the knowledge retention test and performed 30% faster. Giving user feedback is important and can significantly improve the learning outcome compared to traditional approaches.

Vincenzi et al. [24] compared four different learning methodologies: printed material, video tape, text annotations activated by mouse interaction and AR. They performed an immediate and long-term recall test. The aim was to find the best learning paradigm. The participants were all students and had no experience of an aircraft oil pump before. The four groups performed an eight-minute session, learning the terminology, functions and locations of the oil pump parts. After a short three-minute break, participants were asked to answer 10 multiple choice questions and five matching questions. They were asked again after seven days in order to measure how much knowledge they retained after one week. The results indicate that participants who received instruction through AR demonstrated significantly better immediate recall results than those who received print instruction or interactive video instruction. Participants in the AR instructional group also scored significantly higher on the long-term recall test than the video group. However, the AR group failed to show significantly better performance than the video group for immediate recall, and the print and interactive video groups for long-term recall.

Haritos and Macchiarella [25] compared video-based, AR-based and print-based learning approaches. They had 36 participants (all students) performing the two conditions. They were divided into three groups with 12 participants each. Each group performed an eight-minute instructional session, learning about the airplane propeller, its mounting bolts and safety instructions. The aim was to determine how much knowledge participants gained through the different training sessions. An immediate and a long-term recall test after seven days was performed. No results were given in this paper. It is not clear which method is the best for training.

### B. Design Guidelines

Users have high expectations regarding the quality and functionality of an application. In order to meet this requirement, an app designer uses established design guidelines (DG) for their user interface development. We present three general DGs which are well-known in the field of user interface design. The first is the DIN EN ISO 9241-110 [26] which delivers seven heuristics for the user interface design. Further are the principles by Nielsen and Molich [27] and the well-known “8 Golden Rules of Interface Design” [28]. Their similarities are listed below:

- **Simple Design**: An application should have all the required functionalities, not more and not less.
- **Feedback**: Feedback from the system is necessary at every user interaction in order to inform the user about the current action.
- **Preknowledge**: The system should take users preknowledge into account, e.g. everyday life symbolisms.
- **Consistency**: The system should use standards and be consistent throughout the entire application.
- **User Mode**: The system should have different modes (tutorial, beginner, intermediate, advanced) in order to allow advanced users to work faster.
- **User Control**: The system should allow to go back, forward and to the starting point at any time.
- **Help**: The system should recognize incorrect user actions and inform them with a clear understandable solution.
- **Error prevention**: The system should prevent errors e.g. ask the user before acting.
- **Individualization**: The system should be adaptable to personal user demands, e.g. different languages.
- **Help & Documentation**: Help should be available. The system should give easy understandable help at the right time.

These principles allow designing a user interface for standard devices such as smartphones or tablets and can be also used for AR devices. Regenbrecht et al. [30] presented suggestions especially for industrial AR after five years’ experience in 10 different projects which should be noticed when designing AR applications for industrial use cases:

- **Real Data integration**: Most of the demo scenarios work pretty well, but when it comes to the first real scenario trial, systems mostly fail because of the data quantity and complexity. Therefore, the usage of real data from the beginning of the project is crucial for the project’s success.
- **Acceptance**: The initial application should be implemented very carefully with the help of key persons who are accepted among their colleagues and work as close as possible in cooperation with AR researchers. He also prefers an “island setup” beside the production line with single-user and a single-task setting.
- **Simplicity**: According to [26], a simple solution is also recommended instead of an application with the highest level of originality or novelty.
- **Added Value**: Its crucial for the project to get started to outline the return of invest (ROI). If a researcher wants to start a project in the automotive industry, using AR, they have to estimate factors like cost and time savings. All mentioned design guidelines are general for user interface design and can be also used for AR app development. More specific design guidelines for AR-based training for assembly and maintenance task should be noticed. Until now, there are just a few guidelines available.

Webel et al. [29] presented four design recommendations which help to train procedural skills. Procedural skills are the
most important abilities in assembly and maintenance task because it is the human capability to follow a task step-by-step. An AR-based application should rather focus on training these skills than guiding a user through the task such as a navigation system. With a strong guidance, the user can become dependent from the technology and may not be able to repeat the task without any assistance. The following recommendations can help to strengthen the procedural skill:

- **Visual Aids:** Direct visual aids are permanent presented information such as 3D models superimposed on the related real environment. Indirect visual aids are additional information, only presented or available for the user when needed (e.g. text annotations, documentation). This concept allows to adapt information during the learning process. Clear and detailed instructions at the beginning are necessary for the trainee to understand and perform the tasks. During the training, information can be gradually reduced with each assembly cycle. This concept is supported by Fitts and Posner [33] who divided the learning process into three stages: cognitive, associative and autonomous. Different amounts of information are needed to perform the task in each of the learning stages.

- **Mental Model Building:** The mental model of an assembly task describes the internal representation of an entire task. Context information such as progress bars can help creating a mental model. It is also favorable for the mental model building to present pre- and post-conditions of the task [34].

- **Passive Learning:** There should be a part in the training where the trainee is not active and only receives information about the task. This concept can help to gain a global picture of the entire task.

- **Haptic hints:** A vibrotactile bracelet allows to give feedback about current actions and can help to prevent errors. Furthermore, this multimodal approach supports human’s memory and therefore the assembly training.

Another study [31] aimed to find the optimum amount of information to be delivered at the same time for novice user. The optimum amount was found to be from four to five information pieces at the same time.

IV. CURRENT LIMITATIONS

The presented studies in Section III show that AR can help to improve the efficiency and quality of assembly training task but all of them have several limitations which will be summed up in this section.

- **Limitation 1- Participants:** Researchers use mostly students, without any assembly experience, in their assessments. Webel [32] mentioned that cognitive skills such as the procedural memory as well as fine motor skills are necessary to perform an assembly task. AR can help train the procedural skills but not fine motor skills. These skills were acquired through years of experience in the assembly domain. Compared to a real assembly worker with years of assembly experience, students might be much slower and their assembly quality worse, using the same technology. Therefore, the existing evaluation results made with students are subjective.

- **Limitation 2- Assembly Task:** Researcher use simple assembly tasks such as LEGO models for their assembly task. The results, made with those easy assembly tasks are not reliable because a repetition of the measurement in a much more complex environment like an engine assembly line might be much worse or even not possible with the used technology (e.g. projection-based AR).

- **Limitation 3- Comparisons:** Researchers tend to compare their AR-based training systems mostly against paper-based or video-based solutions. Until now, there are no comparisons between an AR-based assembly training and face-to-face training, which is the current training solution in the automotive industry [35].

- **Limitation 4- KPIs:** Researchers mostly focus on measuring the time, which is an insufficient variable for training. There are few evaluations, measuring the quality and the training transfer (immediate recall and long-term recall), which should be favored when evaluating AR-based training systems.

- **Limitation 5- Technology:** Hand-held devices such as tablets and smartphones as well as monitors are used in most of the studies. When assembling, users need to work hands-free, which is not possible with such a device. A possible solution would be a HMD. HMDs are becoming less bulky and lighter. They can be used for assembly training, but until now, there has been no evaluation in an industrial environment. Future studies should investigate the potential of HMDs such as the Microsoft HoloLens for assembly training measuring the usability, the quality, the efficiency of the training, ergonomic aspects (e.g. NASA-TLX) as well as the impact on short- and long-term memory.

V. FUTURE RESEARCH

We indicated five research limitations in Section IV. In Section V, we present our approach to address these limitations and how we want to gain new knowledge. We use the DSR for future research (Fig. 3). This framework is most notable in Computer Science and aims to develop and evaluate new artifacts such as algorithms, new human-computer interfaces or methodologies.
Hevner et al. [36] presented seven guidelines for DSR which will be addressed in our future research. Our aim is to develop and evaluate an HMD-based training for assembly tasks (guideline 1). Werrlich et al. [35] analyzed an assembly training in an industrial environment and found out that the current assembly training is time consuming and stress provoking for new employees. Furthermore, the pedagogical skill varies between different trainers and therefore the training is not standardized. We try to solve these problems with our artifact (guideline 2). In Section III, several design guidelines which we use for the development of our first application were presented. Werrlich et al. [35] did semi-structured interviews and found basic requirements for the development of an HMD-based assembly training. We use that knowledge for our application development and evaluate (guideline 3 and guideline 5) our artifact, regarding the utility and usability, with assembly employees under field conditions in an industrial environment using standardized evaluation techniques (e.g. System Usability Scale). We assess and improve our artifact continuous (guideline 6). Therefore, this research approach provides clear research contributions through the assessments, measuring additionally the quality of the training as well as the short and long-term memory. We will also deliver the first comparison between a HMD-based assembly training (Fig. 2) and a face-to-face training (Fig. 3) and therefore address the fourth DSR guideline.

We present our research contributions to the scientific community as well as to management audiences (guideline 7).

VI. CONCLUSION

In this paper, we gave a comprehensive overview on evaluations and design guidelines for AR-based assembly training tasks. We used a structured research method and searched in four online databases. We used the keywords “augmented reality” AND training and got 862 results. We indicated 17 relevant user studies and seven design guidelines which we will apply for our future application development. Based on our findings, we indicated five scientific limitations such as participants, simple user tasks, lack of comparisons, insufficient KPIs and the technology used in the studies. We will address these limitations in our future work using DSR.

We will develop an HMD-based engine assembly training using the Microsoft HoloLens and evaluate our application in a real industrial environment with skilled assembly workers. Our future evaluations are divided in two phases. The first is the evaluation phase, the formative evaluation, which focuses on iterative designing, assessing and improving our application, while the second phase, the summative evaluation, focuses on the comparison between HMD-based training and traditional face-to-face training.

ACKNOWLEDGMENT

We would like to thank the BMW Group for its assistance and for providing the tools, space and funding to realize and complete the project successfully.

REFERENCES


