2019

university of groningen
faculty of economics and business

Newcastle University
Business School

Intelligent Augmented Reality Assembly Training

Exploring the evolution of assembly training towards smart training methods

Double Degree Master Thesis

N. Heidler (S3903656/180457645)

Rijksuniversiteit Groningen	Newcastle University Business School
Supervisor: Dr Jos Bokhorst	Supervisor: Dr Adrian Small

Master Thesis by Niklas Heidler

Eichenstraße 4, D-58285 Gevelsberg

niklas.heidler@web.de

Intelligent Augmented Reality - Exploring the evolution of assembly training towards smart training methods

December 2019

Word count (excluding references and appendices): 16.165

Double Degree Master

Rijksuniversiteit Groningen
Faculty of Economics and Business (FEB)
MSc. Technology Operations Management
Student No.: S3903656

Newcastle University
Business School
MSc. Supply Chain and Operations Management
Student No.: 180457645

Abstract

Augmented Reality (AR) superimposes information into a real environment and offers a possibility to give information without distracting. In assembly training, much information has to be passed to the trainee theoretically and practically. AR training systems are amongst the most frequently regarded possibilities to support trainers in the assembly training. While a substantial number of such systems have been developed for theoretical situations, only very few companies already implemented them. With proceeding research and technology AR is getting increasingly attractive to companies so that at the moment several are running projects aiming to implement AR into their assembly training.

Therefore, this research aims at empirically giving an overview of which of the different AR systems are best suited for which phase of industrial assembly training and academically exploring the possibilities of the systems. The focus will be laid on non-adaptive AR training systems and more advanced Augmented Reality Adaptive Tutors (ARATs), compared to traditional training methods.

Through conducting a multiple case study in two companies followed by a three-round Delphi investigation with experts in the relevant fields, the current practice of assembly training has been assessed and analysed. Ranking lists have been established of the most important factors of both systems. Furthermore, the Delphi experts came up with a blueprint for an assembly training system using AR.

The results reveal the potential that ARATs will be offering in the future, but also that those are not necessarily beneficial under all circumstances compared to non-adaptive AR training systems. Especially for the last two phases of assembly training, ARATs show a big potential and are, therefore, expected to be a part of the technologisation of industrial assembly.

Key words: Augmented Reality, AR, ARAT, ITS, artificial intelligence, industrial assembly, assembly training, smart manufacturing

Acknowledgements

This thesis marks the end of my MSc degrees in Technology Operations Management and Supply Chain and Operations Management. I am feeling very honoured and proud to receive those degrees from the Rijksuniversiteit Groningen and Newcastle University, two excellent and yet very different universities.

Throughout the past eight months, I have spent many hours exploring an entirely new and future-oriented technology. Still, this research would have been impossible to achieve without any support. I herewith want to thank my supervisors Dr Jos Bokhorst and Dr Adrian Small for always giving me constructive and helpful feedback and spending quite some hours with me discussing the thesis.

A special thanks has to go to Dr Jannes Slomp and his project team researching at the Hogeschool Arnhem Nijmegen (HAN) on augmented reality in assembly for their insights, dedication and contacts to the companies. I hope that my research contributes well to your project.

Furthermore, I want to display my gratefulness to all the experts and practitioners who sacrificed their time for helping me in my research, be it by opening their companies' doors for me, by introducing me to some basic concepts or by answering my questions in the interviews or the Delphi questionnaires. Although some appointments were hard to arrange and I had to drop one case company, every one of you contributed substantially in this research. This is certainly not something one can take for granted.

Lastly, this thesis also finishes my long studying journey, which has brought me to eight different cities. I always had brilliant people around me who all contributed to bring me where I am today. However, the biggest contribution in this journey belongs to my parents who always supported me and made quite some things possible for me that I would have considered impossible.

Groningen, December 2019

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AI .... Artificial Intelligence
AR .... Augmented Reality
ARAT .... Augmented Reality Adaptive Tutor
AV .... Augmented Virtuality
CAD .... Computer-aided Design
CBT .... Computer-based Training
CIP .... Continuous Improvement Process
ELT .... Experiential Learning Theory
HMD .... Head-mounted Display
ITS .... Intelligent Tutoring Systems
JI .... Job Instruction
JM .... Job Methods
JR .... Job Relations
KPI .... Key Performance Indicator
MR .... Mixed Reality
RV .... Reality-Virtuality
TWI .... Training Within Industries
UI .... User Interface
VR .... Virtual Reality

1 Introduction

Skilled assembly workers are the backbone of well-functioning assembly operations. However, due to the lack of qualified personnel (Coad *et al.*, 2016; Ramteke, 2019), many companies are struggling to effectively train their workers without negatively affecting the operations. Therefore, they are exploring new ways to effectively and efficiently train new employees.

Augmented Reality (AR) was perceived as science fiction for decades until it had its mainstream breakthrough when 21 million gamers used an AR smartphone game within one week after the release (Serino et al., 2016). The technology however does not only entertain, but already found recognition earlier in manufacturing research (Caudell and Mizell, 1992). In recent years, the advancement of computer technologies motivated researchers to develop new ideas and uses for AR. Amongst the most frequently regarded ones is the usage of AR to train industrial assembly tasks (e.g. Gavish et al., 2015; Westerfield et al., 2015; Werrlich et al., 2018) as this could help to solve the struggles to train workers while maintaining the same output quality (BMW Group, 2019).

The training of assembly skills is a process that stayed unchanged for decades as methodologies like the Training Within Industries (TWI) Job Instruction (JI) extensively proved their effectivity and efficiency (Dinero, 2005). Nowadays however people who could train others new skills are rare. As companies are aiming to utilise their qualified personnel as efficiently as possible, the high utilization in training is getting increasingly problematic. Through AR in assembly training, the personnel need for assembly training can get reduced (BMW Group, 2019).

Most AR assembly training publications focus on software development and do not take training methods sufficiently into consideration. However, most of the more practically oriented systems incorporate many aspects from existing training theory (e.g. Herbert *et al.*, 2018; Werrlich *et al.*, 2018; BMW Group, 2019) which implies that industry is not intending to reinvent the wheel of assembly training.

Although there is already a considerable amount of different types of AR-based training systems (superimposed objects (Xu *et al.*, 2008; Li *et al.*, 2009; Hořejší, 2015; Werrlich *et al.*, 2018), superimposed instructions (Kreft *et al.*, 2009; Webel *et al.*, 2012), force feedback systems (Charoenseang and Panjan, 2011)) and they proved to be functional in their intended area (amongst others Morkos *et al.*, 2012; Webel *et al.*, 2012; Hořejší, 2015) still no agreement could be reached about how to approach the development of those systems.

This uncertainty regarding the design of AR-based assembly training systems increased even more when Westerfield et al. (2015) combined AR technology with artificial intelligence (AI) based intelligent tutoring systems (ITS) to create an adaptive system based on situation evaluation. As such Augmented Reality Adaptive Tutors (ARATs) recognize the operator's actions, they enable the provision of feedback and, therefore, the coverage of a wider range of training approaches, methods and phases (Herbert *et al.*, 2018).

This leaves companies willing to implement AR-based assembly training systems two kinds of functioning systems, but little guidance on which to choose and how to design systems them (Herbert *et al.*, 2018).

Regarding effectiveness, several developers found that their non-adaptive AR training systems were more effective than traditional, non-technological training methods (Li et al., 2009; Hořejší, 2015; Werrlich et al., 2018). Even 25% more effective than those non-adaptive systems is the advanced ITS-based Motherboard Assembly Tutor Westerfield *et al.* (2015) came up with. Herbert et al. (2018) further proved the enhanced effectivity (see further section 2.4). However, the BMW Group (2019) recently decided to implemented non-adaptive AR in their engine assembly training (Günnel, 2019). So although the ARAT systems are better performing, some companies still prefer non-adaptive AR. Therefore, this thesis aims at exploring where and under which circumstances which AR system is preferable.

Consequently, the question for which phases of the industrial assembly training process which kind of AR training system (non-adaptive or ARAT) offers more value is addressed. To answer this question the current practice of industrial assembly training will be observed to establish the practical reasons companies are choosing AR for their training systems. Furthermore, the benefits and pitfalls of ARATs in assisting the training process will be assessed and the extra requirements that ARATs have compared to non-adaptive AR training systems will be listed and ranked.

This research will contribute to the body of knowledge on AR assembly training systems by giving insights about which phases of industrial assembly training to assist with which kind of AR training system most effectively. It provides companies with a decision guideline on whether and which kind of AR training system to choose, based on the nature of the phases where assistance is desired. Lastly, it gives assembly characteristics supporting the implementation of an ARAT.

The remainder of this thesis is structured as follows: The second chapter will review the existing literature on training methods, industrial assembly training, AR-based training systems and AR assembly training systems to provide a sound background on those topics. After that, the research questions for this research will be formulated based on the literature review. Subsequently, the research methodology will be described. The fifth chapter describes the results, which are discussed in the sixth chapter. Finally, the research is concluded in the seventh chapter.

2 Literature Review

In the following section, the relevant literature will be reviewed. This has the purpose to get to know the most influential learning theories and assess to which extent they are put into action in assembly training, AR training systems and AR-based assembly training systems. Furthermore, an overview of the existing training systems will be provided. This chapter is separated into the sections of training systems, industrial assembly practices, AR training and AR assembly training.

2.1 Training systems

Training systems are systematic approaches to how an operator can achieve new skills and capabilities (Wang et al., 2016). Training has to be distinguished from guidance where workers receive instructions on already known processes as part of a continuous improvement process (CIP) (Haagsman, 2017). Although not the focus of this research, some insights from training might also be useful in guidance conditions.

The aspect of systematic approaches to train and educate employees has long attracted attention in the scientific fields of behavioural psychology. Two of the most relevant learning and training theories, the experiential learning theory (ELT) and TWI JI are regarded in this section before in the last part the technological training solution of ITS will be reviewed.

2.1.1 Experiential Learning Theory (ELT)

One of the most frequently regarded training theories is the ELT, initially formulated by Dewey (1938). The idea behind this is that people learn best if they experience the relationships and consequences of their action directly, like e.g. in the well-known trial and error heuristic (Jueptner *et al.*, 1997; Kolb *et al.*, 2001). Kolb (1984) identified and described the three main models of ELT, out of which two are relevant for this work: (1) The Lewinian Model of Action Research and Laboratory Training and (2) Dewey's Model of Learning. Piaget's Model of Learning and Cognitive Development focuses on the development of learning from childhood to adult and is, therefore, not further described.

Lewin (1951) saw experiential learning as a four-stage cycle of concrete experiences causing observations and reflections. Those then lead to the formation of abstract concepts and generalizations which are tested in the following and again lead to an experience. This concept emphasizes the importance of feedback processes in acquiring new skills.

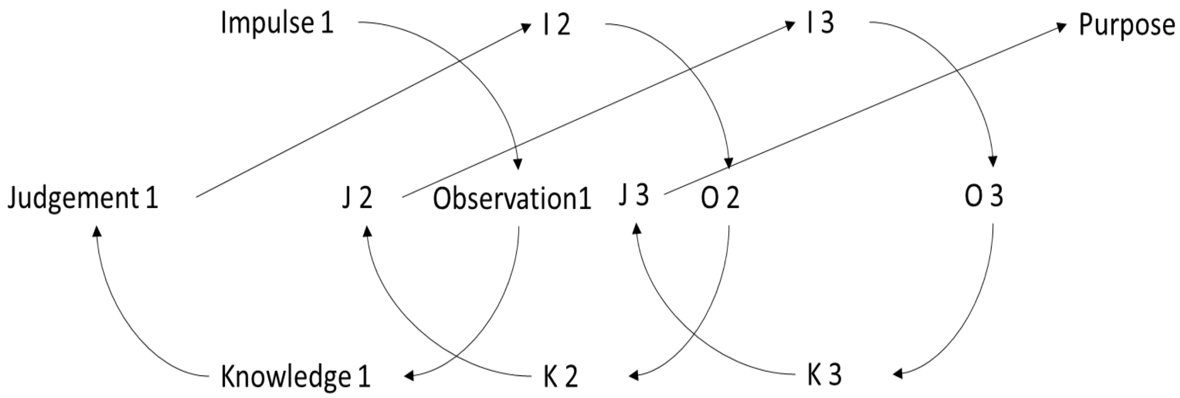

FIGURE 2.1: DEWEY'S MODEL OF EXPERIENTIAL LEARNING (KOLB, 1984)

Dewey (1938) stresses even more the importance of obtaining feedback in order to learn new skills as his four-stage cycle of impulse, observation, knowledge and judgement explicitly leads to a new impulse (see Figure 2.1). As this cycle continues several times, the learners acquire incrementally their new skill(s).

Those concepts have been summarized by Kolb (1984) via six basic theses about ELT:

1) *"Learning is best conceived as a process, not in terms of outcomes. […]*
2) *Learning is a continuous process grounded in experience. […]*
3) *The process of learning requires the resolution of conflicts between dialectically opposed modes of adaptation to the world. […]*
4) *Learning is a holistic process of adaptation to the world. […]*
5) *Learning involves transactions between the person and the environment. […]*
6) *Learning is the process of creating knowledge*."

Dewey (1938, p. 25) captured the essence of those theses in one sentence: "Any experience is mis-educative that has the effect of arresting or distorting the growth of further experience." The idea of ELT is to educate through erasing the barriers missing experiences build and avoiding the experiences where this could not be done. Therefore, a mixture between guiding the trainees and letting them try out is the key for successful learning.

To conclude, ELT provides some psychological insights into how humans acquire new knowledge. Although it initially differed substantially from the idealist approaches of traditional, more theoretical education with help of books and was seen sceptically, it proved to be very influential (Kolb, 1984) and got connected to the development of more adaptive AR training systems by Herbert *et al.* (2018) (see section 2.4.2).

2.1.2 Training within Industry (TWI)

TWI was originally a US governmental department built for quickly scaling up the American Industry for World War II. Later, Dooley (1945) summarized the measures and analysed the results as the programs deployed offered also potential in a non-war economy. While most Western economies focused on further optimizing industrial mass production after the war, the Japanese soon started to form Lean Management based on the essentials of TWI (Dinero, 2005). With the increasing spread of Lean, TWI grew in importance as it supports the implementation of Lean in an employee-centred manner (ibid.).

The TWI method is composed of the four modules of Job Instruction (JI), Job Methods (JM), Job Relations (JR) and Program Development. While JM focuses on continuous improvement, JR on solving conflicts between employees and program development on the identification of processes to be improved, JI established guidelines how to teach new skills effectively (Dooley, 1945). Therefore, all references to TWI in the remainder of this thesis will be referring to JI as this is the only relevant module regarding this research.

JI gives a general four-step training method summarized for the teachers on JI cards to take with them:

> *"Step 1 – prepare the worker*
>
> > *Put him at ease.*
> >
> > *State the job and find out what he already knows about it.*
> >
> > *Get him interested in learning job.*
> >
> > *Place in correct position*
>
> *Step 2 – Present the operation*
>
> > *Tell, show and illustrate one IMPORTANT STEP at a time.*
> >
> > *Stress each KEY POINT.*
> >
> > *Instruct clearly, completely, and patiently, but no more than he can master.*
>
> *Step 3 – try-out performance*
>
> > *Have him do the job – correct errors.*
> >
> > *Have him explain each KEY POINT to you as he does the job again.*
> >
> > *Make sure he understands.*
> >
> > *Continue until YOU know HE knows.*
>
> *Step 4 – follow-up*
>
> > *Put him on his own. Designate to whom he goes for help.*
> >
> > *Check frequently. Encourage questions.*
> >
> > *Taper off extra coaching and close follow-up."* (Dooley, 1945, pp. 158-159)

Even though the aforementioned steps might seem outdated nowadays regarding the development assembly took since its development, the fact that TWI has been a central part of Toyota's employee training for over 50 years and is increasingly used in other companies emphasizes the importance and timelessness of these methods as they still determine how an employee should be trained efficiently (Dinero, 2005; Werrlich et al., 2018). Therefore, TWI is still an immensely relevant training method that numerous companies follow.

2.1.3 Intelligent Tutoring Systems (ITS)

ITSs are using AI to adaptively teach skills depending on the prior knowledge (Sleeman and Brown, 1982; as cited by Dermeval *et al.*, 2018). They are a subset of the group of computer-based training (CBT) systems, so systems where computers deliver instructions to the trainees (Alqahtani and Ramzan, 2019). They are defined there by their intelligent adaptability in real-time scenarios generating a response "as close to a human response as possible" (Alqahtani and Ramzan, 2019, p. 14) and the computer-based imitation of human tutoring based on a "one-on-one dialogue […] helping the student learn something" (Evens and Michael, 2006, p. 3).

Early versions of ITS like the ELM-ART (Brusilovsky et al., 1996; Weber and Brusilovsky, 2001) or INSPIRE (Grigoriadou et al., 2001) were online intelligent interactive integrated textbooks adapting to the learners' knowledge level and learning style. Modern systems go further and include psychological and technological aspects, which yielded a drastic improvement in effectiveness (Alqahtani and Ramzan, 2019). Such systems like the VALERIE (Petrovica and Ekenel, 2016) or Gnu-Tutor (Ivanova, 2013) use video cameras, microphones, physiological sensors or eye trackers to recognize the trainees' emotions and adapt to them. This data is used to refine the model, provide insights about the student to the teacher (if still present) and giving signals and feedback to the student (directly or via teacher). A possible interaction system with an ITS and a tutor both responsible for tutoring collaboratively is illustrated in Figure 2.2. In this system, the human tutor is only responsible for intervening if the

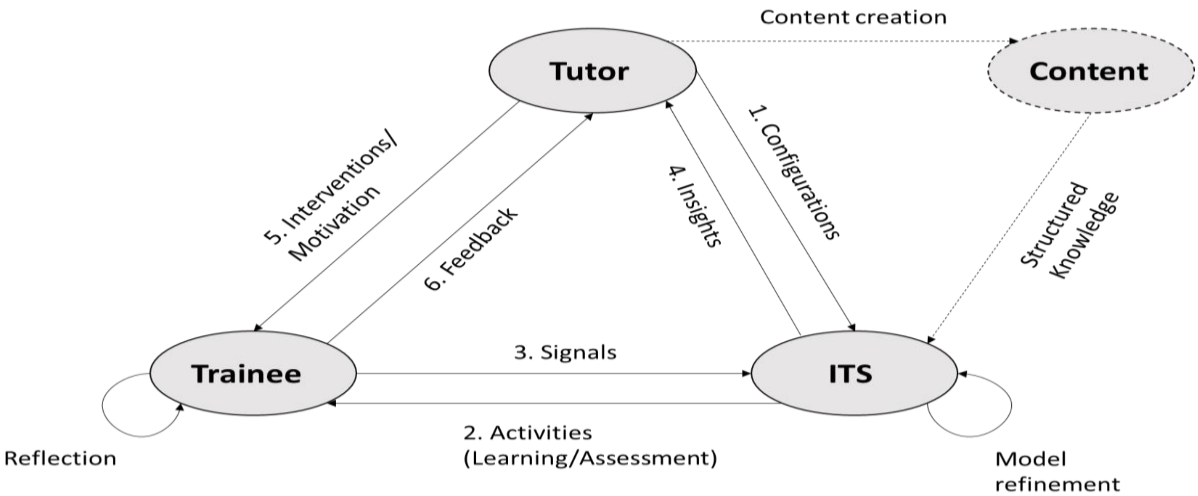

FIGURE 2.2: INTERACTIONS IN AN ITS WITH A TUTOR (ADAPTED FROM KOKKU *ET AL.*, 2018)

student does not get the system's instructions, motivating the student and configuring and monitoring the system.

Another focus of ITS research is the design of authoring tools for passing over the design to non-programmer authors (Dermeval et al., 2018). Those systems like the ASPIRE by Mitrovic et al. (2009) or the Mathtutor by Aleven et al. (2009) open up the highly complicated ITSs to the domain experts – teachers and trainers of diverse subjects. One of those systems, the xPST from Gilbert et al. (2015) also included an interface to 3D game engines, which are frequently used as the basis for AR systems (see section 2.3.3).

Alqahtani and Ramzan (2019) state the main goals of modern ITSs as (1) creating an user interface (UI) enhancing the visibility of useful data, (2) individually assessing the individual steps in the learning process, (3) providing context-specific hints and explanations and (4) presenting a problem tailored for the user. The mixed reality (MR) environment AR creates offers a strong opportunity for the first two (see section 2.3.1). The third goal is one of the main strengths of ITSs while the fourth is hardly possible to realize in the standardized assembly environment this research regards.

ITSs are often stated as the second most efficient training method after human tutoring (Kulik and Fletcher, 2016; Alqahtani and Ramzan, 2019). However, VanLehn et al. (2007) even found that tutoring through ITS was as effective for Novice students learning intermediate-level content as a human tutor connected via a chatroom. Therefore, it can be summarized that ITSs are a worthy and effective alternative to human tutoring especially their scalability offers strong perspectives regarding cost-efficiency (Kokku et al., 2018).

Regarding assembly training possibilities, ITSs have the potential to expand the technological assistance in the learning process compared to current processes as their adaptability allows a more dynamic training than the static training e.g. video tutorials offer. Furthermore, this adaptability entails the possibility to provide the trainee learning experiences. They could, therefore, be a key technology to enhance the adaption of ELT in practice, with the main weakness that looking at a desktop PC distracts from the assembly task (Herbert *et al.*, 2018).

2.2 Industrial assembly practices

Assembly processes are central in several manufacturing industries (Al-Ahmari *et al.*, 2018). In the following section, some assembly characteristics will be described which might influence the applicability of different training systems. Subsequently, it is analysed how assembly workers are trained in practice.

2.2.1 Assembly characteristics

Assembly is defined as "aggregation of all processes by which various parts and subassemblies are built together to form a complete […] assembly or product" (Nof *et al.*, 2012, p. 2). It is classified into manual assembly, special purpose assembly and flexible, programmable assembly. While the two latter rely mostly on automated machines, manual assembly involves a worker that has to be trained (Nof *et al.*, 2012).

Within this broad definition, assemblies can be categorized among different characteristics. Ranz *et al.* (2018) found five different characteristics groups in which different attributes could be taken by the assembly. Out of the five groups, the factors of economics, product traits and processes will be further regarded as the system inclusion and the safety features are a given for the assembly training processes.

From the economic perspective, the tact time is regarded as the most important factor for the training. When the tact time dictates a fast assembly to the workers, the trainees have to be prepared more thoroughly to keep up to that.

The group of product traits entails the attributes of weight, stability, manipulability, sensitivity and value. All those are important factors determining how easy it is to handle the product due to either physical or economic reasons (Hammerstingl and Reinhart, 2017).

Regarding the process, the product variance and the required accuracy are relevant factors for assembly training. While the variance influences how broad the training has to be, the accuracy has a direct influence on the experience required to perform the job.

A frequently regarded assembly characteristic not covered by Ranz *et al.* (2018) is the assembly complexity. Haagsman (2017) assessed assembly complexity via 17 distinctions of factors in four categories. Falck *et al.* (2017) distinguish assembly complexity considering time dependence (static or dynamic) and origins (basic or perceived). The latter distinction is seen as more beneficial for this research as it gives a simple and straightforward concept of complexity.

The relevant assembly characteristics are summarised in Table 2.1.

Characteristics	Attributes
Complexity	Static or dynamic Basic or perceived
Economics	Tact time
Product traits	Weight Stability Sensitivity Manipulability Value
Process	Product variance Required accuracy

Table 2.1: Relevant assembly characteristics (adapted from Falck *et al.*, 2017; Ranz *et al.*, 2018)

As to the author's knowledge no research exists on how most of those assembly characteristics influence the choice of training methods, it is not possible to develop valid hypotheses on the influences of specific factors. Therefore, most of them will be left uncontrolled in this research and the results will be observed.

However, a meta-analysis of different developed training approaches revealed a possible interaction between the product value and the virtuality (compare Figure 2.5 in section 2.3.1) of the approach. The comparison of 18 different papers developing assembly training systems revealed that the more valuable a product, the more virtual the training method used is (see Figure 2.3). Researchers frequently chose fully non-physical virtual reality (VR) approaches for valuable products like plane parts or medical products (e.g. Xia *et al.*, 2012; Ho *et al.*, 2018; Nash *et al.*, 2018) while the assembly of less valuable products or substitutes is often based on experiential learning (e.g. Pozzi *et al.*, 2014; De Vin and Jacobsson, 2017; Ahmad *et al.*, 2018) allowing more mistakes (see section 2.1.1). Between those, there are AR training systems allowing the trainees to experience the assembly physically while avoiding costly mistakes (Werrlich *et al.*, 2018). The products here are mostly products of medium value, like motherboards (Westerfield *et al.*, 2015), gully traps (Hořejší, 2015), actuators (Webel *et al.*, 2012; Gavish *et al.*, 2015) or water pumps (Boud *et al.*, 1999).

Therefore, it is expected that companies working on implementing AR into their assembly training are mostly assembling medium valuable products (ca. 500 € to 8000 €).

Overall, there is a great variety of assembly characteristics and most of them are assumed to influence the optimal training methodology. However, there is no research investigating this. Therefore, this

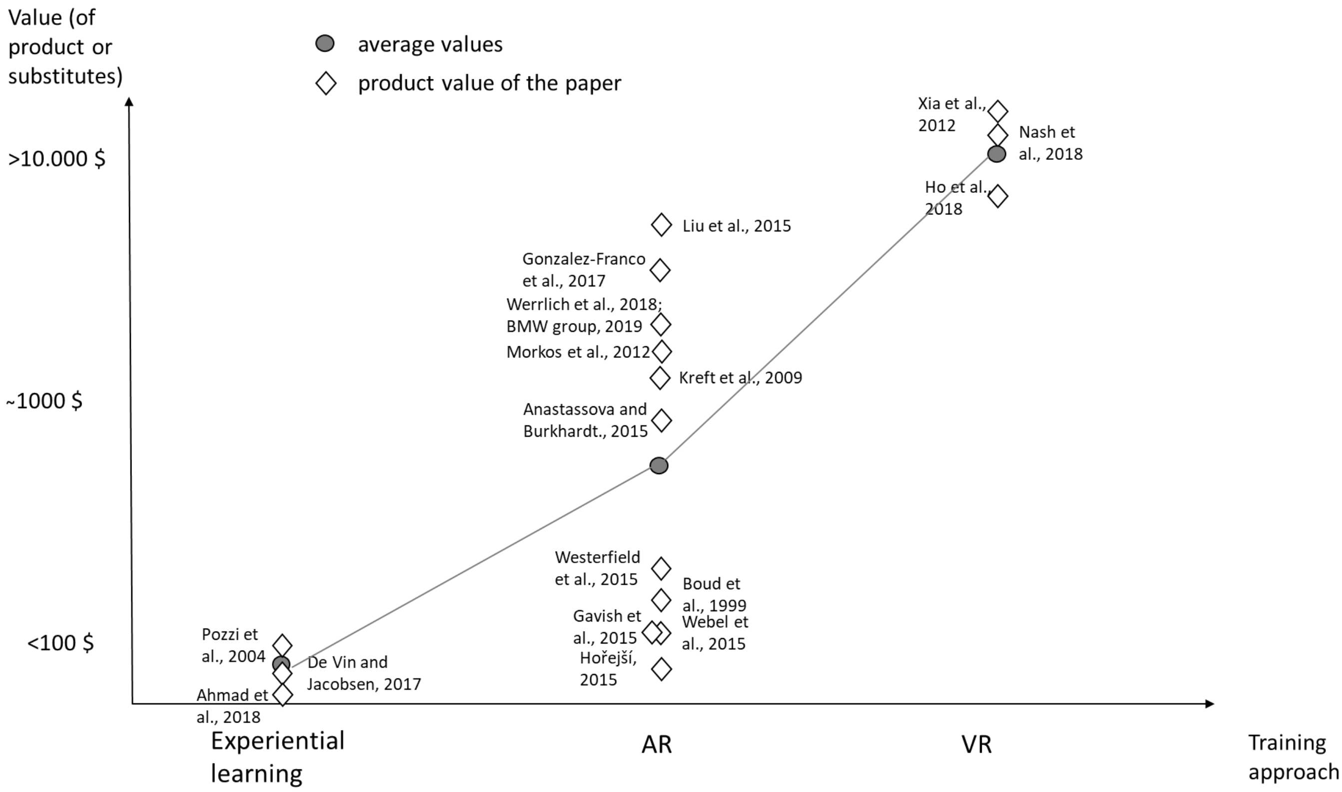

FIGURE 2.3: POSITIONING OF DIFFERENT TRAINING APPROACHES REGARDING THE ESTIMATED PRODUCT VALUE

research focuses on the one aspect for which evidence was found through a meta-analysis, the product value. The aim is to practically verify the observations made. The influence of the remaining aspects might be assessed by future research.

2.2.2 Assembly training

Assembly training is the process of providing employees with the skills to assemble the product or part without errors in an effective, time-efficient manner (Nöhring *et al.*, 2015). While tutoring is the "one-on-one dialogue[…] between a teacher and a student to help the student learn something" (Evens and Michael, 2006, p. 3) training entails all processes involved in the skill acquisition. Werrlich *et al.* (2018) define the aims of assembly training to "acquire procedural as well as fine-motor skills" (p.463).

Nöhring *et al.* (2015) distinguish between passive training methods like lectures and active training involving the students. The active methods are further divided into compiling methods encouraging the student's activity and explorative methods giving great independence and responsibility to the students. While only 5% of knowledge taught by passive training methods is retained, up to 75% are for active methods (Brauer, 2014; as cited by Nöhring *et al.*, 2015). This supports the implications of the ELT (see section 2.1.1) that learning needs practical experiences. Although those numbers speak in favour of active learning, passive learning is still effective for basic information and introductions to a topic (Nöhring *et al.*, 2015).

The input for the active training part in assembly training can be given in several different ways. While traditional methods are based on demonstrations and paper instructions, several digital and virtual methods grew importance using technologies like additive manufacturing or VR (Langley *et al.*, 2016; Al-Ahmari *et al.*, 2018).

Abele *et al.* (2017) see the current actions during assembly training in the continuous delivery of engineering competencies and a strong multidisciplinary education and training background. However, Cachay *et al.* (2012) found that those methods have limitations and that action-oriented learning events yield a greater application-performance and a higher degree of action-substantiation knowledge. Therefore, Abele *et al.* (2017) conclude that new learning approaches in manufacturing need to allow training in realistic manufacturing environments.

The Learning Factory concept addresses these requirements. Learning Factories are defined by the CIRP encyclopaedia (Abele, 2016) as a learning environment with authentic processes, a changeable setting resembling a value chain where a physical product is manufactured while following a didactical concept (see Figure 2.4). The broader definition also includes virtual value chains, service products and remote learning (ibid.).

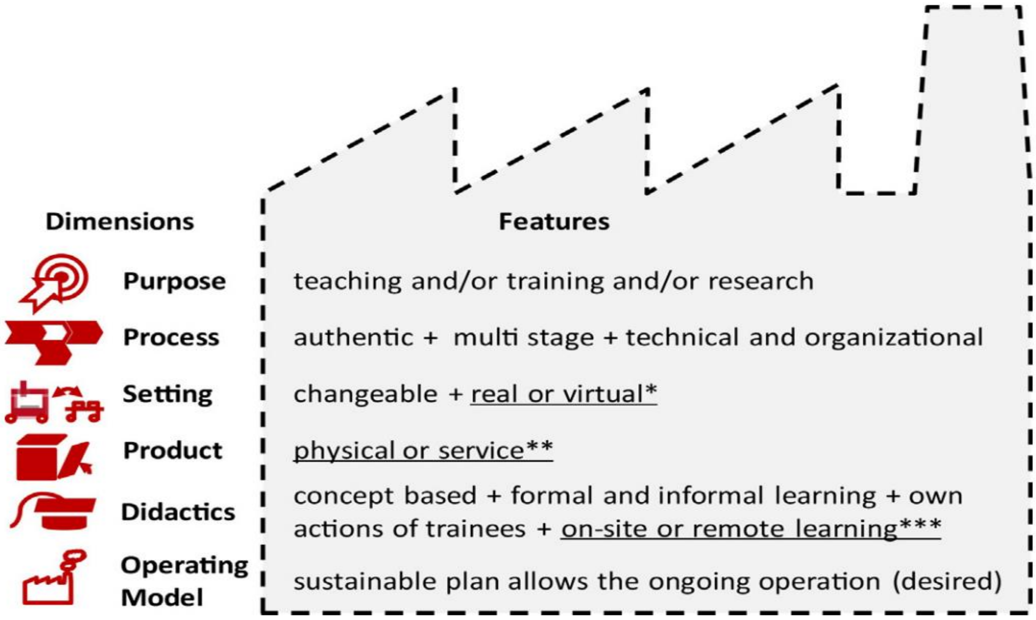

FIGURE 2.4: KEY CHARACTERISTICS OF LEARNING FACTORIES (ABELE ET AL., 2017)

Major benefits of the Learning Factory are the experiential learning possibilities where trainees have the freedom to create own implementations and test them (see section 2.1.1) and the inclusion of didactical concepts (Abele *et al.*, 2017). Regarding the continuing technological advancements, the didactical concept, operating model and process to be followed could be delivered by an ITS (see section 2.1.3).

Assembly training is overall an important field where the digitalization offers interesting development perspectives. Evidence shows that active training methods are more effective, but are also more costly, mainly due to personnel costs. New technologies like additive manufacturing, VR or AR have the potential to weaken that trade-off and, therefore, enhance effective and efficient training, especially in an environment like a Learning Factory where the training takes place in a realistic educational setting. The integration of ITSs in such a system could offer further benefits.

2.3 AR training

AR is defined as the technology set enabling the user to “see the real world, with objects superimposed upon or composited with the real world” (Azuma, 1997, p. 356). Although the potential of AR training is already known for more than two decades in research (Azuma, 1997), the topic is recently gaining relevance thanks to fast technological advancements resulting in first systems being used in practice, like at the BMW Group (2019).

In the following, AR training will be classified and the software and hardware used will be described.

2.3.1 AR training classification

FIGURE 2.5: SIMPLIFIED REPRESENTATION OF A REALITY-VIRTUALITY CONTINUUM (MILGRAM *ET AL.*, 1995)

The exact classification of AR, especially relative to MR is still disputed in research. On the one hand, Milgram et al. (1994) classified AR at the reality-virtuality (RV) continuum as an integrative central part of MR, but closer to the real environment compared to augmented virtuality (AV) which involves real-world elements in a virtual environment. VR is the other end of the spectrum and describes the user being in an entirely virtual environment (see Figure 2.5). On the other hand, Yamamoto (1999) established MR as a mixture of AR and AV, so the middle of the continuum. In this research, the definition of Milgram et al. (1994) will be adapted as it is the most accepted and straightforward. AR is, therefore, seen as a part of MR with a higher degree of real elements complemented by superimposed virtual elements (see Figure 2.5).

Boud et al. (1999) established the distinction between context-free systems as static images being superimposed by virtual images and context-aware systems continually adapting the virtual environment to the reality it is facing.

While the former distinction on the reality-virtuality continuum is still relevant in literature (Ternier et al., 2012; Neges et al., 2018), the latter lost relevance as the emergence of technology-enabled better context-aware systems (Daponte et al., 2014) making the context-free ones obsolete.

Among the first ones, Kaufmann (2004) developed an AR training system and brought it to practice in geometry classes. The positive impression of this motivated several others to develop AR-based training systems in different educational areas, such as language skills (Liu, 2009; Jee et al., 2011) or spatial abilities (Dünser et al., 2006; Martín-Gutiérrez et al., 2010). Lee (2012), Billinghurst and Dünser (2012) and Santos et al. (2014) all evaluated the different systems considering their usability in education with varying results. This emphasizes the importance of prototype testing for usability and benefits in the learning process (Santos et al., 2014).

However, the systems are all highly specified on their specific environment and, therefore, only have limited generalizability. As an answer to the missing possibilities to generalize the systems, authoring tools were designed allowing non-programmers to design an AR learning system (e.g. Lucrecia et al., 2013; Jee et al., 2014). Leblanc et al. (2010), furthermore, acknowledged that a combination of technological and traditional training yields better learning results speaking against fully technological training methods.

The two main components of AR training systems are software and hardware. Software components need to enable the developers to quickly program and enable the usage of AR in its environment. Furthermore, they should allow different components the integration into an AR system, e.g. through an authoring tool (Ong et al., 2008). Hardware systems meanwhile have to make the software run smoothly and give the users the desired input without negatively influencing their performance through e.g. big weight or unpractical cable connections.

In the following two sections, the state of the art of those central parts of AR systems will be analysed. Afterwards, the knowledge on AR systems in general will be carried over to analyse applications of AR systems in assembly training.

2.3.2 Hardware

The hardware forms the visible part of any AR system. Its task is to transfer the content into the augmented environment in the least distracting manner (Azuma, 1997).

While Milgram et al. (1995) classified AR display solutions broadly as either see-through or monitor based displays, Syberfeldt et al. (2016) categorized the modern solutions differently and identified video-based glasses, optical glasses, a video-based tablet or a spatial projector as possibilities for AR hardware. The optimal choice amongst those depends on the circumstances of usage (Syberfeldt *et al.*, 2016).

In general, most AR researchers utilise off-the-shelf optical glasses respectively head-mounted displays (HMDs) as they are seen as the most flexible and easy to operate (Azuma, 1997; Ong et al., 2008; Novak-Marcincin et al., 2013). However, for AR educational systems Santos et al. (2014) found an even distribution between desktop monitors, handheld devices, an overhead projector and HMDs as hardware choices.

Accordingly, also the hardware used in AR assembly training systems is more diverse, also regarding the nature of the augmentation. The majority of researchers use visual hardware as augmentation solutions. Most systems run with HMDs (Boud *et al.*, 1999; Morkos *et al.*, 2012; Westerfield *et al.*, 2015; Syberfeldt *et al.*, 2016; Werrlich *et al.*, 2018; BMW Group, 2019; Ferrati *et al.*, 2019) or similar video-based glasses (Li et al., 2009; Charoenseang and Panjan, 2011) as those leave the operator’s hands free to handle the product. Nevertheless, some authors see advantages in tablets as solutions due to their higher robustness (Webel et al., 2012; Gavish et al., 2015). For stationary workstations, ordinary computers are seen as the best solution because of their higher computing power. They are either combined with a monitor (Hořejší, 2015; Liu et al., 2015a; Liu et al., 2015b) or a projector (Mura *et al.*, 2016). However, the occupancy of at least one hand while handling the PC and the distraction from the assembly product is a significant disadvantage in the environment of assembly operations (Herbert *et al.*, 2018). Nee *et al.* (2012) also see potential to use projectors for portable solutions in AR assembly training, but to the knowledge of the researcher, no such system has already been developed.

Some advancements also included devices to address further human senses in the reality augmentation. As an assembly environment entails many physical aspects, mostly the sense of feeling was included by bracelets (Webel et al., 2012; Gavish et al., 2015) or an exoskeleton (Charoenseang and Panjan, 2011). The sense of hearing was used for audio feedback or instructions by Boud et al. (1999), Kreft et al. (2009), Aouam *et al.* (2018) and Ferrati *et al.* (2019).

For improving the dataset offered to the software to acknowledge the environment and the actions by the operator, Kreft et al. (2009) and Charoenseang and Panjan (2011) included force sensors on the product or integrated in a pair of gloves in the hardware system.

2.3.3 Software

While the majority of researchers build a system with off the shelf hardware, the availability of ready-to-use software is very limited. Generally, two types of AR software have to be distinguished. First, there is software developed specifically for the one system it is used in (e.g. Mura et al., 2016; Danielsson et al., 2017). Second, there are authoring tools aimed at providing non-programmers with an interface to develop their own AR system within specific areas (e.g. Lucrecia et al., 2013; Jee et al., 2014). Due to the early stage of the research and the variety of researched applications a big variance within the categories regarding functionality and software interface has established.

Despite serving different purposes, all software applications have in common that the researchers developed them themselves, although often based on standard software interfaces. While some software sets like “Unity-3D” (Danielsson et al., 2017; Tatić and Tešić, 2017; Werrlich et al., 2018) or “Unifeye SDK Metaio” (Hořejší, 2015; Mura et al., 2016) were used more frequently, the variety of different software solutions is generally very high. Furthermore, many research papers do not describe which software tools were used (e.g. Li *et al.*, 2009; Webel *et al.*, 2012; Gavish *et al.*, 2015).

The functionality offered by the software applications differs significantly due to their different purposes. Generally, AR training software mostly superimposes work instructions (e.g. Tatić and Tešić, 2017), objects (e.g. Jee et al., 2011; Aouam et al., 2018) or virtual human-resembling tutors (Jee et al., 2011) into the real environment.

Considering the used learning approaches, only Kaufmann (2004) explicitly state the theories considered. Nonetheless, most authors propose systems consistent with important learning theories like ELT, e.g. by enabling students to see and experience geometrical forms (Dünser *et al.*, 2006; Martín-Gutiérrez *et al.*, 2010).

However, as most systems are limited to giving information in a predefined order they lack responsiveness to the operator's actions and performance. Therefore, the aspects of feedback provision and deviation from the optimal path to experience the consequences which are central in TWI (see section 2.1.2) and ELT (see section 2.1.1) are not covered by those systems.

2.4 AR assembly training

Throughout the years, the potential for AR in assembly training got increasingly acknowledged in literature. After some first acknowledgements of a potential use of the technology in assembly training and guidance (Caudell and Mizell, 1992; Azuma, 1997) the first systems were developed and tested in the late 1990s. By now, a considerable amount of systems has been designed. The BMW Group (2019) was the first major company to implement AR into their assembly employee training, resulting in a change from one-to-one teaching to three-to-one teaching meaning that only one tutor now reaches the same outcome quality with three trainees that he/she had with one before.

The performance of AR assembly training systems can be measured in terms of efficiency and effectiveness. In literature, effectiveness is covered through the measurements of outcome quality, so the number of mistakes made (Li *et al.*, 2009; Webel *et al.*, 2012; Gavish *et al.*, 2015; Westerfield *et al.*, 2015; Werrlich *et al.*, 2018) or assembly speed (Boud *et al.*, 1999; Huenerfauth, 2014; Hořejší, 2015; Westerfield *et al.*, 2015), sometimes displayed as a learning curve (Hořejší, 2015). As measures of system efficiency, the number of trainers needed per trainee to yield a certain outcome (Morkos *et al.*, 2012; BMW Group, 2019), throughput times in the training (Morkos *et al.*, 2012) or cost estimations to run and implement such a system (Anastassova and Burkhardt, 2009; Kreft *et al.*, 2009; Li *et al.*, 2009; Huenerfauth, 2014) are most frequently used.

In the following section, the status quo of AR systems in assembly training is described by introducing the two categories of non-adaptive AR systems and ARATs.

2.4.1 Non-adaptive systems

As non-adaptive training AR systems, in this research all systems are considered that do not offer the automatic adaptation of the system to external factors. Therefore, this section mostly focuses on the differences in the non-adaptive software.

As described in section 2.3.2, the hardware used most often off-the-shelf and rather similar to each other. Regarding the software however, the fact that by now nearly all systems have been self-developed indicates some variety, although the differences are relatively small.

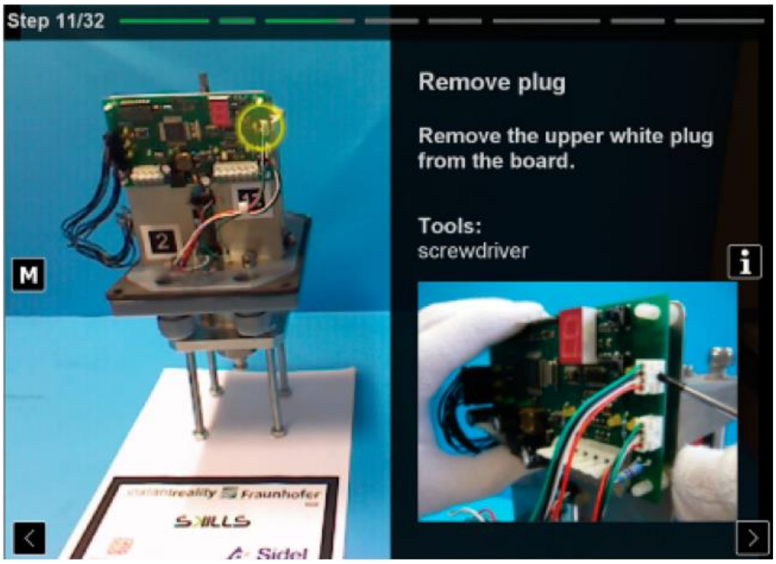

FIGURE 2.6: IMAGE OF THE INDUSTRIAL MAINTENANCE AND ASSEMBLY-AR SYSTEM (GAVISH *ET AL.*, 2015)

Considering the software functionality, the spectrum goes from superimposing work instructions and labels in the real image (Hořejší, 2015; Mura et al., 2016; Syberfeldt et al., 2016; Tatić and Tešić, 2017) via integrating 3D parts from the computer-aided design (CAD) file in the image (Hou and Wang, 2012; Hou *et al.*, 2013; Rentzos *et al.*, 2013; Aouam *et al.*, 2018) to combinations of both (Kreft *et al.*, 2009; Morkos *et al.*, 2012; Webel *et al.*, 2012; Gavish *et al.*, 2015; Hořejší, 2015; Liu *et al.*, 2015b; Danielsson *et al.*, 2017; Ferrati *et al.*, 2019). Regarding which functionality would be suited best under which conditions, no clear indication was found.

The educational outputs of the systems are mostly relatively basic. Even though most researchers combine 3D objects and work instructions in their systems, the systems still offer only limited learning benefits as they are very standardized and lack personalization for the operator (see e.g. Figure 2.6). A summary table of the different applications of AR in assembly is provided in Appendix A.

Herbert et al. (2018) acknowledged the linear nature of the instructions, the lack of knowledge checks and the scaffolded instruction layout as major limitations of such non-adaptive AR training systems.

One of the main motivations of researchers to focus on AR systems in assembly training is the hope to reduce the necessity of human trainers. This can be seen explicitly in the publication of the BMW Group (2019) and implicitly in several other publications (Gavish *et al.*, 2015; Danielsson *et al.*, 2017; Tatić and Tešić, 2017; Cohen *et al.*, 2018; Ferrati *et al.*, 2019).

In order to assess the results the systems achieve, comparisons to ordinary training methods are frequently used. In those comparisons most researchers found better values for outcome quality (Xu *et al.*, 2008; Li *et al.*, 2009; Webel *et al.*, 2012; Gavish *et al.*, 2015; Werrlich *et al.*, 2018; Ferrati *et al.*, 2019) and assembly speed (Boud *et al.*, 1999; Huenerfauth, 2014; Hořejší, 2015; Cohen *et al.*, 2018; Ferrati *et al.*, 2019) of the product. However, Werrlich *et al.* (2018) and Gavish *et al.* (2015) found that learning times were slower for AR training, which they explained with the fact that most users were not used to the AR technology and its handling.

Noticeably only few researchers explicitly base their systems on the training approaches of ELT (see Herbert *et al.*, 2018) or TWI (Werrlich et al., 2018). Most developed training systems base the learning journey on the paradigm that people already learn from just getting instructions and assembly information (e.g. Li *et al.*, 2009; Hořejší, 2015). However, ELT (see section 2.1.1) and TWI (see section 2.1.2) both emphasize the importance of letting the trainees have failures and providing them feedback on their performance. This needs soft- and hardware that recognizes and adapts to the operators' actions like ITSs are capable of (see section 2.1.3). Herbert *et al.* (2018) identified this as a research gap to be covered by future research.

2.4.2 Augmented Reality Adaptive Tutors (ARATs)

As reaction to the low educational value most non-adaptive AR assembly training systems deliver, Westerfield et al. (2015) combined AR with ITSs in a design space which Herbert et al. (2018) defined as ARATs to enhance the learning via more intelligent software. The self-evaluation of Westerfield *et al.* (2015) and the considerations by Herbert et al. (2018) reveal promising and yet mostly unexplored possibilities with the usage of ARAT systems offering enhanced usability, intuitive conveying, accurate mental model development and environmentally shaped experiences.

The ideal ARAT system envisioned by Herbert *et al.* (2018) combines augmented environments from AR technology with psychomotor learning through adaptive ITSs. In psychomotor learning, the trainees are "using motor skills and precision in physical tasks to integrate domain knowledge" (Herbert *et al.*, 2018, p. 166), so they learn through experiencing as proposed by the ELT (see section 2.1.1). In this way, the system displays feedback via an AR device on the performed actions (see Figure 2.7) and automatically displays the next instruction when the preceding one has been performed. Therefore, the operator is left with more freedom to experience himself what deviations from the optimum lead to and the handling of the system is simplified. As a comparison, in non-adaptive AR systems the operator manually goes through the static steps without getting any feedback from the system.

FIGURE 2.7: IMAGE OF THE MOTHERBOARD ASSEMBLY TUTOR (WESTERFIELD *ET AL.*, 2015)

This computational operation evaluating the past actions could be either integrated in the AR software (intelligent client-based ARAT) or in the ITS software (non-intelligent client-based ARAT).This means that developers can decide between using an already existing ITS and designing an interface to the AR software or modifying the ITS to integrate it into the AR software (Herbert *et al.*, 2018).

The computational power to run such advanced systems could be provided by wearable computing systems (Kreft *et al.*, 2009). The bigger amount of hardware enables more complex computations as needed for running AI and AR at the same time smoothly.

The automated adaptation of the software-generated tutoring on the operator's abilities opens new possibilities for AR assembly training systems, as they could then be enabled to assist the operator thoroughly in the steps 2 to 4 of the TWI system (see section 2.1.2). Therefore, the inclusion of AI-based systems like ITSs in AR training systems is offering a big potential which shall be explored within this research and compared to non-adaptive AR systems in assembly training.

3 Research Questions

As established earlier, even though there is considerable interest in researching AR training in manufacturing settings and assembly, it seems apparent that there is a lack of structured, general approaches how to implement an AR training system effectively. Some general considerations of AR have been published to identify potential in AR systems (Azuma, 1997) or identify key performance indicators (KPIs) and beneficiaries (Jetter et al., 2018). However, none of the publications regarded for what phases in assembly training AR can assist, regardless of which kind. Furthermore, only Herbert et al. (2018) and partly Huenerfauth (2014) analysed AR assembly training systems globally not focusing on only one system. In light of the inclusion of ITS in AR training systems (Westerfield *et al.*, 2015; Herbert *et al.*, 2018), the question arises for which assembly training phases a more advanced ARAT can assist while a non-adaptive AR training system cannot do so.

Therefore, the main research question of this thesis is the following:

> *RQ: Which phases in the industrial assembly training could be assisted best by which kind of AR training system?*

In order to answer that question, at first the phases of industrial assembly training have to be observed. Through the observation, insights will be gained on what the process looks like in practice, where it differs from theory and which system capabilities are needed so that the training follows the same procedure established as optimal and the trainees obtain the skills they need. This is done under the assumption that the case companies' current training system is optimal for their circumstances. Furthermore, the AR assembly training systems they are currently developing will be assessed.

After that, the different kinds of systems have to be compared regarding their possibilities to assist in the established assembly training phases. This is done by working out the main advantages and disadvantages that the systems have in assembly training. A special focus will be laid on reasons to choose AR in assembly training in the first place and subsequently assessing the benefits ARAT systems have compared to non-adaptive AR training systems.

Finally, the decision between a non-adaptive AR-based assembly training system and an ARAT is a trade-off decision between the higher effectiveness of ARATs in the training and the higher requirements to program, implement and run those. As those extra requirements entail higher costs, they have to be assessed to provide companies with a decision aid on which system might offer the most value-added in their specific settings.

Consequently, the research question is backed up by the following sub-questions:

SQ1: What are the phases of industrial assembly training in practice?

SQ2: Where does practice of industrial assembly training differ from theory?

SQ3: What do the current AR assembly systems of companies look like?

SQ4: What are the practical advantages of non-adaptive AR training motivating companies to invest into the implementation of such a system instead of relying solely on human tutor-based learning? What are the disadvantages speaking against an investment? What do the practical systems look like?

SQ5: What are the practical advantages that might motivate companies to invest into the implementation of an ARAT system rather than non-adaptive AR? Which disadvantages speak against this?

SQ6: What additional requirements compared to non-adaptive AR training systems do ARATs have to be implemented, run and maintained?

By answering those research questions, a considerable contribution to the body of research can be achieved as the existing and newly generated knowledge is aggregated in a decision aid for companies thinking about what kind of AR assembly training system to implement.

4 Research Design

After deriving the research questions, the ways how to answer them have to be identified. Therefore, in the following section the methodology will be described. Furthermore, the ethical issues that had to be regarded during the research will be briefly discussed. Finally, the measures taken to ensure the reliability and validity of the gained data are summarized.

4.1 Methodology

This research was conducted on a pragmatic philosophy because its aims were of a practical nature (Saunders *et al.*, 2019, p. 145). As the main aim was to discern and generate a pattern for which tasks in assembly training AR technology is useful, the research nature is inductive (Karlsson, 2016, p. 21).

Regarding the sub-questions, it becomes apparent that SQ1, SQ2 and SQ3 are descriptive, SQ4 and SQ5 are explanatory and SQ6 is exploratory. The main RQ is of inductive exploratory nature which indicates the necessity for open research approaches like expert interviews, focus groups or literature searches (Karlsson, 2016, p. 21; Saunders *et al.*, 2019, pp. 186-188).

Due to the width of the range the research question and its sub-questions cover, a mixed methods research design was chosen, with a multiple case study and a Delphi method (see Table 4.1). This mixed design is regarded as providing the best opportunities to combine different approaches for diverse aspects into one outcome based on real-life information (Sekaran and Bougie, 2016; Saunders *et al.*, 2019). Specifically, this means that the practical corporate descriptive data to answer mostly SQ1, SQ2 and SQ3 was gathered through the appropriate measure of the case study (Yin, 2014). The explanatory SQ4 and SQ5 and the exploratory SQ6 were mostly covered by the Delphi method and only to a small extent by the case study as the combination of knowledge of different aspects helps to form new, yet unknown knowledge and identify issues in a field that is still to be explored (Schmidt, 1997; Laick, 2012). As the Delphi method is built on knowledge gained by the case study, the research was sequentially designed (Sekaran and Bougie, 2016).

Research design aspect	Choice
Unit of analysis	Assembly training
Independent variable	Type of AR support
Dependent variable	Training efficiency and effectiveness
Research philosophy	Pragmatism
Approach	Inductive
Research strategy	Mixed methods research
Methodological mixture	Interviews, Observations and Delphi method
Time horizon	Cross-sectional

Table 4.1: Summary of research methodology

This research methodology aimed at first finding out in the case study how assembly training is conducted nowadays without AR, so to establish the status quo of assembly training. Based on this knowledge, the Delphi method had the target of exploring how far this status quo could be assisted or taken over by the different kinds of AR technology.

The following sections will describe the used methods thoroughly and motivate their aims within this research.

4.1.1 Multiple Case Study

The first part of the research was a multiple case study in two companies. As Meredith (1998) points out, a close examination of few cases through a multiple case study enables a researcher to build new theories which can be generalized afterwards. Therefore, the case study aimed at getting an impression of practical assembly training at the two case companies.

The case studies were carried out following the case study protocol by Yin (2014, pp. 84-85), which can be found in Appendix B. The cases were selected according to a theoretical replication logic, so in hope of contrary results due to predictable reasons (Voss et al., 2016). The differences amongst the cases were desired in the production volume and related to this in the organisation of the assembly, so the allocation of tasks between different employees and machines. The main expectation here was to find different levels of standardization in the assembly and, therefore, also in the assembly training methods. This would then also influence the possible uses of AR in assembly training. The assembly characteristics defined in section 2.2.1 were intentionally left uncontrolled and only observed. Another criterion for the case companies was the necessity of interest to implement AR technologies in their assembly training as this ensures the relevance of this research in the investigated field. The interaction of product value and level of virtuality of the training approach assessed in section 2.2.1 justifies the expectation of similar medium range values for those companies.

The selected case companies, therefore, represented a cross-section of different production volumes and ways to organise the assembly with a similar range of product value (see Table 4.2). A third case company was contacted, but the researcher could not acquire all the information needed from them.

Throughout the cross-sectional case study, semi-structured interviews have been conducted with trainers and trainees and the training session of one case company could be observed in detail (Sekaran and Bougie, 2016). With those methods, it was aimed

Characteristic	Case Company 1	Case Company 2
Product value	100-10.000 €	900-10.000 €
Production volume	Low	High
Assembly organisation	Manual assembly with workstations	Assembly line
Trainers	Team Leaders and foremen	Experienced line workers with extra training
Trainees attributes	Disadvantaged and disabled	Untrained; no disabilities

TABLE 4.2: CASE COMPANY CHARACTERISTICS

mainly to collect data for SQ1, SQ2 and SQ3, with the semi-structured interviews providing data for SQ1 to SQ6.

Inter-viewee	Case Company	Position
1.1	1	Supervisor of product assembly, project leader of the AR project
1.2	1	Assembly foremen and trainer
2.3	2	Production Manager and Project Leader of the AR project
2.4	2	Assembly line worker and trainer
2.5	2	Assembly line trainee
0.6	0	Management Trainee Inclusive Fieldlab for technological solutions

TABLE 4.3: LIST OF INTERVIEWEES

In every case company, multiple people were interviewed in order to get an unbiased and complete picture of their training methods and their implementation plans for AR in assembly training (Karlsson, 2016). The interviewees were selected using a theoretical replication logic regarding their perspective on the assembly training processes. The perspectives from people theoretically constructing the guidelines for training and people participating in the training as trainer or trainee gave a holistic picture of the processes. Interviewee 0.6 forms an exception as no further perspectives could have been gained from her company. Therefore the company did not serve as case company and was coded as case company 0, but the inputs given by the interviewee could still be analysed. A list of the interviewees is provided in Table 4.3.

The results of the case study were analysed regarding the research questions and the expected outcomes and compared with the other case in a cross-case analysis.

4.1.2 Delphi Method

After the case study led to some insights about the practice of industrial assembly training, a three-round Delphi method was conducted to combine opinions of experts in the different relevant fields on the two kinds of AR training systems first without interference and afterwards with taking into account the opinions of their peers (Laick, 2012). The methodology of the Delphi investigation was captured in a Delphi protocol (see Appendix C).

The experts have been selected based on their knowledge in one or several relevant research fields and the organisational requirements of willingness and time availability to participate in the defined periods (Adler and Ziglio, 1996; as cited by Laick, 2012). Overall, 15 possible experts have been identified and contacted. Due to absences and time constraints, the final panel consisted of ten experts. The participation among those amounted to 83% across all three rounds (see Table 4.4), with seven responses in the first and nine in the second and third round.

The three round's design followed the three-phase design of Schmidt (1997), which has been frequently used in ranking Delphi questionnaires (e.g. Hasson *et al.*, 2000; Schmidt *et al.*, 2001; Okoli

Expert	Expertise	Expertise origin	Participation
1	AR	Developing AR software	3/3
2	AR	Researching & Consulting on AR solutions	2/3
3	AR	Researching & Consulting on AR solutions	2/3
4	AI	Researching AI	3/3
5	AI	Researching AI	3/3
6	AI	Researching AI	2/3
7	Training Methods	Working as Training Consultant	3/3
8	Assembly Training	Production coordinator and group member of AR training project	1/3
9	Assembly Training	Supervisor of product assembly, project leader of the AR project at case company 1	3/3
10	Assembly Training	Production Manager and Project Leader of the AR project at case company 2	3/3

TABLE 4.4: DELPHI METHOD PARTICIPANTS

and Pawlowski, 2004). In this design, the expert panel was first asked to brainstorm about a topic without restrictions. The mentioned aspects then got validated and narrowed down in the second phase before in the third phase the experts ranked the importance of them (see Figure 4.1).

Specifically, the results of the case study were put into a first questionnaire openly asking for important aspects of ordinary AR training systems and ARATs in assembly training and comparisons of those (see Appendix E1). The content of those answers was then grouped and given to the experts in the second round. This round served two objectives: The experts validated the grouped answers and also narrowed down those lists to only the aspects the majority rated as important (see Appendix E2). In the third round, the experts then ranked the aspects among their importance (see Appendix E3). The rankings were then aggregated calculating the mean rank given and Kendall's coefficient of concordance w (Siegel, 1988, pp. 262-271). This is a coefficient taking values between 0 and 1. The higher the value, the higher is the consensus between the given ranks (ibid.). The threshold

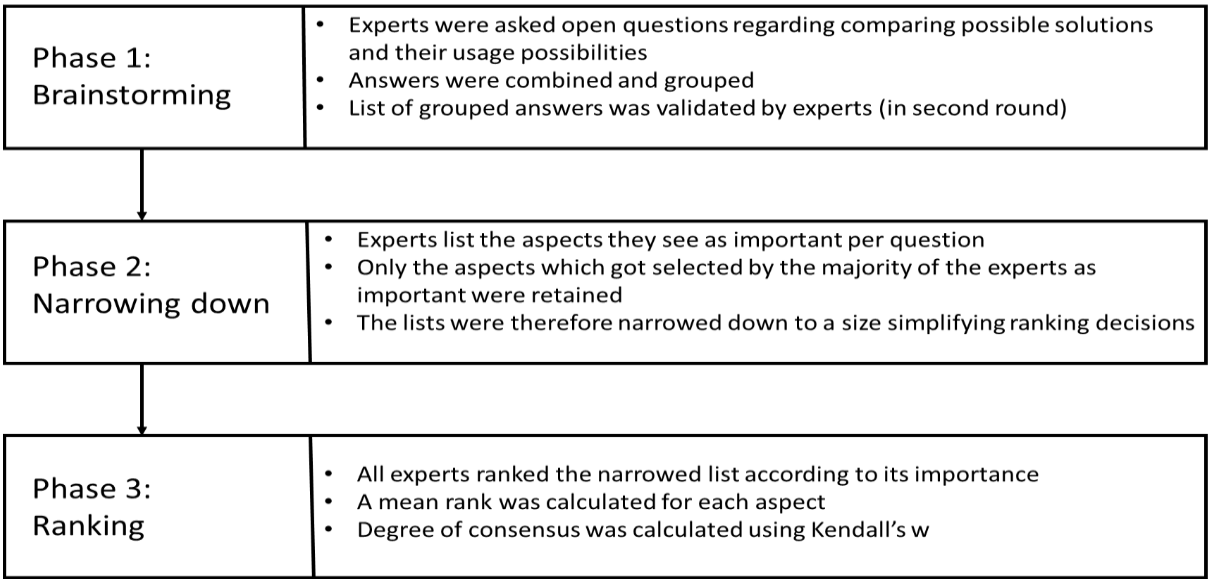

FIGURE 4.1: DELPHI INVESTIGATION PROCESS (ADAPTED FROM SCHMIDT *ET AL.*, 2001)

for the coefficient indicating a statistically significant consensus differs depending on the number of entities to be ranked (N) and the number of judges (k) and can be taken from a table Siegel (1988) gave. The formula for Kendall's w is provided in Appendix F.

The ranked lists were the final results. All intermediary results in form of aspects given and their vote from round 2 are given in Appendix G. The questionnaires have been sent out electronically via 'Google Forms' and the participants were given one week to answer them. They were sent out in CW 40, 42 and 43 of 2019.

The aim of the Delphi method was to establish a consensus among the combined expertise of the participants in the numerous different relevant fields. Through the questionnaire and the briefing on the preceding round's responses each questionnaire contained, the experts could not only give their own opinion but also had the opportunity to learn about other experts' opinions and incorporate those in their own perceptions. The emergence of the opinions was studied with this method as well as the content and expertise the participants offered to the researcher (Laick, 2012).

4.2 Ethical considerations

For ensuring valid results, it critically important that ethical guidelines considering all stakeholders of the research are followed. Therefore, the ethical issues at the different stages of the thesis progress as defined by Saunders *et al.* (2019) will be discussed in the following.

4.2.1 Ethical issues during design and gaining access

Gaining access to data and insights from practitioners is the first critical step for conducting a research ethically.

As this research is part of an umbrella project by the Hogeschool Arnhem Nijmegen (HAN) to explore the possibilities of AR in assembly, the experts of companies involved in this project could have been feeling a pressure to participate in the research. To avoid this, all participants were approached without any coercion or inducements at acceptable risks associated and were given the possibility to withdraw from the study at any time. Through fully informing via participant information sheets, an informed consent was gained from all participants (Saunders *et al.*, 2019).

Special attention in this phase required the observation of current training methods during which pictures were taken. This practice was agreed to in advance by the company under the condition that no faces would be visible on the pictures.

Furthermore, all interviewees got informed on the research procedure and their rights. Everyone signalled their agreement to those through signing a consent form in advance.

4.2.2 Ethical issues during data collection

During the data collection, several critical issues to be handled ethically correct were endangered to occur.

The participants of the case study or the Delphi method might have changed their mind about participating. For leaving them the freedom to do so, a possibility to withdraw from the study was given and introduced before starting the study and coercion was put on the participants at any time (Saunders *et al.*, 2019).

Furthermore, the data collection period is crucial for keeping the research objective. Therefore, all actions were documented so that the research is fully replicable and the reader can assess the objectivity himself.

As the participants kept confidentially, all actions were taken to ensure this including keeping their data only on secured University servers, always contacting them personally (especially for E-Mails) and anonymizing them in the report (Saunders *et al.*, 2019).

Additionally, the observation raised the issue of reactivity, so the reactions of people to being observed (Bryman and Bell, 2007, p. 139). However, as neither a covert study nor a habituation of the observed people would have been applicable due to practical and time constraints, the researcher always stayed in the background to keep the distraction and disturbance minimal.

Lastly, also the researcher’s security had to be ensured during the research (Saunders *et al.*, 2019). This was done a risk assessment before starting the research and a general caution towards all possible harms.

4.2.3 Ethical issues related to analysis and reporting

During the final research stage when the data is analysed and reported, the subjects of objectivity and confidentiality continued to be the main issue.

Special attention was granted here towards maintaining the confidentiality of the organizations and participants in a way that it is also not possible to identify them indirectly by information given in the report like age, gender, position in the company and the company’s professional field (Saunders *et al.*, 2019). This issue was handled by only giving the data necessary to understand the research in the report. The participant’s personal and contact data needed for the research were kept confidential at secured University servers at all times and will be deleted after submitting the thesis.

Furthermore, it had to be ensured that in the analysis of the cases and the Delphi rounds no information was misrepresented, falsely excluded or not objectively analysed. This was done by two validation rounds during the research.

4.3 Reliability and Validity

To ensure the quality of scientific research and its implications, some measures have to be taken to achieve reliability and validity of the research, especially in case study research (Bryman and Bell, 2007; Hair, 2007; Voss et al., 2016). Therefore, the measures taken to realize a high-quality research are summarized in Table 4.5.

Requirement	Measures to achieve requirement
Reliability	• Utilization of a research protocol ensuring repeatability and neutrality of research (Yin, 2014) • The questionnaires in the Delphi method were designed in a way that internal consistency could be assessed, increasing the reliability (Hair, 2007) • All steps of the research are documented to establish a coherent chain of evidence (see e.g. Appendix D) • Participants validated interview transcriptions and answer groupings
External validity	• Multiple case studies protect against observer bias and are more likely to deliver good results (Voss et al., 2016) • Usage of theoretical replication logic ensured neutral & constant replications (Voss et al., 2016)
Construct validity	• Multiple evidence sources and case companies were used to get a holistic impression • The documentation of the research ensures a complete and coherent chain of evidence that the reader can retrace
Internal validity	• Multiple respondents per case company reduced subjectivity in the research

TABLE 4.5: MEASURES TO ENSURE RESEARCH QUALITY (BASED ON VOSS ET AL., 2016)

5 Results

In the following section, the results of the research are described and analysed scientifically before they will be discussed later.

The aim of this research was to find out the possibilities of current non-adaptive AR and the future ARAT technology in the setting of industrial assembly training. Therefore, this study was asking which phases in the industrial assembly training could be assisted best by which kind of AR system. In the following, the answer to this question will be given separated into the different time horizons investigated.

5.1 Status quo of AR in industrial assembly training

In the following sections, the current status of AR in assembly training practice will be described (SQ1 – SQ3). This was researched by the case study only.

5.1.1 Practice of industrial assembly training (SQ1) and its gap to theory (SQ2)

Regarding the practical assembly training today, it turned out that both companies follow training methodologies slowly building up the worker's capabilities (see Table 5.1). Case company 1 uses for this purpose dummy orders which they designed to "assemble and disassemble constantly for training purposes" (Interviewee 1.2) to teach the basic skills while case company 2 directly starts to train on the job. Furthermore, both companies work with high levels of interaction between trainers and trainees. All Interviewees emphasized that interaction between trainer and trainee is crucial for the training and both companies assign one trainer for one trainee to achieve a good and personalized training.

It became visible that the skilled trainers were during the training less productive than they would normally be. However, case company 2 managed to let the trainers' job be increasingly taken over by the trainees on the running assembly line. Only occasionally the trainers had to intervene by doing the tasks of the trainees very quickly to keep up with the assembly line when the trainees were a bit too slow. In contrast, case company 1, with their more

Case company 1	**Case company 2**
Done step by step: 1. Train basic skills with dummy orders 2. Demonstrate 3. Let trainee do it under assistance 4. Slowly reduce monitoring when he/she knows the task	Done step by step: 1. Company introduction 2. Demonstrate it 3. Explain important things 4. Let the trainee do it, check and answer question

TABLE 5.1: CASE COMPANY TRAINING METHODOLOGIES

manual assemblies, reported a considerably lower productivity of their employees while they were teaching the trainees which mostly happened off the job.

Looking at it from a broader perspective, it becomes apparent that both case companies follow training methodologies close to TWI, with minor differences between each other. In case company 1, the phases 2, 3 and 4 can be related to the steps 2, 3 and 4 of the TWI methodology. The training methodology of case company 2 relates in all four phases to the steps of TWI (compare section 2.1.2) on this broad perspective.

In detail however, the approaches differ significantly between each other. While case company 2 followed the TWI methodology in a fixed and standardized manner to get the employee fully prepared for the standardized work at the assembly line, case company 1 lays their focus more on personal adaptation and close monitoring of the trainees during and after the training. This not done with the intention to give the trainees more possibilities to try out parts of the assembly, but rather because their employees need much supervision due to their disabilities.

Looking at the standard TWI methodology, it becomes apparent that case company 1 leaves out the first step of preparation and instead assesses the new trainees during the training on dummy orders. Case company 2 meanwhile prepares its trainees for the job by introducing them to company, location and security aspects rather than introducing them to their new job as proposed by TWI. Both companies also deviate from TWI in step 4 with not yet leaving the trainees alone, but keeping them under a decreasingly close supervision with the trainer still having full responsibility. As all observed differences between TWI and the companies' training methodologies are located in the first or last step, it can be summarised that the steps 2 and 3 form the basic skeleton of TWI. In the end, both companies have their "own interpretation" (interviewee 2.3) of the TWI methodology, but still follow it in its main aspects.

The ELT, however, did not play a role in the companies. Both companies justified this with striving for perfection so that trainees cannot be allowed to learn by their own experience a possibly non-perfect way. Herbert *et al.* (2018) expected ELT to play a role in assembly training and based several statements on ARATs on that assumption. Although these cannot be considered as falsified by this case study only, the results raise doubts in this.

Overall it can be summarized that corporate practice in training is in most aspects as described in theory, but only in TWI theory. It was reported by several case study participants that the two central steps 2 and 3 of the TWI methodology form the basic skeleton for their training methodology which is then adapted to the specialities of the company's assembly.

5.1.2 Practical systems under development (SQ3)

Regarding the systems currently under development, the two case companies investigated are both working on bespoke-tailored software solutions to run their future non-adaptive AR training systems on off-the-shelf HMDs. As they need mobility and free hands to operate the product for the trainees, a stationary solution would not work for them. The desired functionalities are still to be defined in detail. To relate these systems to the assembly characteristics defined in section 2.2.1, the characteristics observed are summarised in Table 5.2. It becomes apparent that the differences between the companies lay mainly in the complexity and product traits.

Characteristics	Case Company 1	Case Company 2
Complexity	*Static*: system well-structured, some products complex to assemble → medium *Dynamic:* high operational unpredictability → high *Basic:* low to medium *Perceived:* most workers are cognitively weak → high	*Static*: system well structured, assembly mostly easy → medium *Dynamic*: low operational unpredictability → low *Basic*: complex products, but tasks mostly simple → medium *Perceived*: short time for tasks, but little knowledge needed → medium
Tact time	No tact time given	Varying between 1 and 4 minutes
Product traits	*Weight:* mostly light, few heavy products *Stability:* mostly high *Sensitivity:* mostly low *Manipulability:* good, few exceptions *Value:* 100-10.000 €	*Weight*: heavy (ca. 70-100 kg) *Stability*: high *Sensitivity*: medium *Manipulability*: low *Value:* 900-10.000 €
Process	*Product variance:* High *Required accuracy:* High	*Product variance:* Low *Required accuracy:* High

TABLE 5.2: OBSERVED ASSEMBLY CHARACTERISTICS OF THE CASE COMPANIES

Interviewee 0.6 reported that her company is developing a smart projector, so a fixed workstation working with an AR projector and AI. Due to the more powerful non-mobile computers available, they are incorporating adaptability similar to what ARATs offer. It is desired to have a fully self-contained system adapting to the trainee's capabilities. Also in this system, the hardware is off-the-shelf and the software is bespoke-tailored. Due to missing information, it cannot be clearly stated whether or not this could be considered an ARAT.

5.1.3 Summary current practice

The case study revealed that practical training methods are relatively close to theory, but only to TWI theory. While both companies' training methods only have minor adjustments to the TWI in the outer steps, ELT as another important training theory did not play a role.

The TWI methodology can be seen as the most promising training method in assembly training and, therefore, the steps of TWI are considered in the remainder as phases of assembly training. The goal of this thesis is then to assess which steps of TWI could be best assisted with which kind of AR technology. Furthermore, as the differences are only minor and depend on the company's preferences, it was decided to not assess in the Delphi investigation the applicability of the AR systems on the adaptions the companies made, but stick to the standard methodology to achieve a better generalisability.

Finally, the systems under development in the companies emphasize the practical relevance of the systemic choice as both types already play a role in practice.

5.2 Future of AR in industrial assembly training

The aforementioned insights about the current industrial assembly training were subsequently incorporated in the three-stage Delphi investigation (see section 4.1.2) such that the participants assessed the applicability of the different AR systems specifically on the TWI steps. As theory implicated that also ELT would play a role, the answers and ideas given were analysed with special attention to whether they display characteristics of ELT.

In the following section, the results regarding the future of AR in assembly training will be described. In each of the rankings given, the lowest average rank stands for the highest assigned importance. The answers from the first two rounds that led the way to the rankings are given in Appendix G. The descriptions are separated regarding the SQs they are covering.

5.2.1 Comparison of non-adaptive AR to ordinary training methods (SQ4)

Table 5.3 shows the ranking of the most important advantages and disadvantages of non-adaptive AR as established by the Delphi participants. As the α=0,05 significance threshold of 0,241 for the advantages was reached, it may be stated that the experts' rankings there are not independent from each other (Siegel, 1988, p. 270), so there is a significant consensus. The critical value of 0,287 for the disadvantages was not reached. However, the higher number of advantages that were seen as important by the majority of respondents in the second round provides evidence for a positive attitude towards non-adaptive AR in assembly training. The missing significance can be seen as a hint that the different disadvantages are similarly important.

Amongst the factors that have not been considered important by the majority of the experts, the factors of better possibilities to spot flaws in educational material and the more visual and interactive level of the training were the two most acknowledged advantages. For the disadvantages, those are the possibly problematic acceptance and the need for trainees to motivate themselves more (see Table A.3 in Appendix G).

Advantages for AR	Average rank	Disadvantages for AR	Average rank
Scalability of the training	2,22	Trainees do not receive personal feedback	2,33
Instruction possibility for many different assembly combinations without the need for extra explanations by the trainer	2,78	Training is not personalized	2,44
Standardization of the training procedures	3,33	Difficult, time-consuming and expensive set-up and development for the training	2,56
Bigger flexibility regarding time, place and speed	3,78	Missing flexibility for small situation changes	2,67
Trainers have more available resources to allow time for other responsibilities	4,00		
High efficiency regarding human resources	4,89		
Degree of consensus (Kendall's w)	**0,254* > crit. value (0,241)**	**Degree of consensus (Kendall's w)**	**0,012 < crit. value (0,287)**

TABLE 5.3: COMPARISON OF NON-ADAPTIVE AR TRAINING TO ORDINARY TRAINING

Additionally, the interviewees from the case study gave the concern that acceptance among the trainees might be a problematic aspect, especially with older trainees not used to modern technology. Interviewee 0.6, furthermore, gave the point that the reduction of human interaction and high screen times for the workers might be causing stress or even illnesses like Epilepsy. Most Delphi experts, however, did not regard those aspects as important.

Amongst all mentioned aspects regarding non-adaptive AR training systems only the more visual perspective on training can be related to ELT, indicating a weak support for ELT by non-adaptive systems.

5.2.2 Comparison of ARATs to non-adaptive AR training (SQ5)

Table 5.4 illustrates the ranking of the most important advantages and disadvantages of ARATs compared to non-adaptive AR training systems. In this comparison, no ranking achieved an $\alpha=0,05$ significant consensus, so the experts' rankings are independent from each other. However, the number of aspects carried on from the second to the third round already displays a positive attitude of the experts towards ARATs. Furthermore, it is visible that ARATs are seen as advantageous for the support of practical training procedures, but disadvantageous in the system management.

The most important advantages of ARATs not ranked in the final round were the possibility for full monitoring and a higher motivation for the trainees through the system. As disadvantage, the aspect

Advantages of ARATs	Average rank	Disadvantages of ARATs	Average rank
Personalized training	2,44	More complex in development and maintenance	1,44
Learning could be done "Just-in-Time", so you receive information right when you need it	2,89	High initial costs	2,11
Adaptability of instructions	3,00	Big datasets required	2,44
Improved training outcome through higher efficiency	3,11		
Systems could gather experiences to improve the teaching	3,56		
Degree of consensus (Kendall's w)	**0,064 < crit. value (0,259)**	**Degree of consensus (Kendall's w)**	**0,259 < crit. value (0,333)**

TABLE 5.4: COMPARISON ARATS TO NON-ADAPTIVE AR TRAINING

that a high reliability of the spatial-awareness would be required for a stable system ended up closest to being ranked in the third round (see Table A.4 in Appendix G).

Similarly to the Delphi experts, the case study interviewees emphasized the interactivity and adaptability of the system as advantages that would motivate them to prefer an ARAT over a non-adaptive AR system.

No evidence was found within this comparison which would support the relevance of ELT.

5.2.3 Additional requirements of ARATs compared to non-adaptive AR training (SQ6)

Requirement	Average rank
Evaluation algorithms for all steps	2,00
Definitions of when, where and how feedback shall be given	2,33
More sensors to create situation-awareness	2,56
Very big data sets regarding the operator's past and present performance and knowledge	3,11
Degree of consensus (Kendall's w)	**0,131 < crit. value (0,287)**

TABLE 5.5: ADDITIONAL REQUIREMENTS OF ARATS COMPARED TO NON-ADAPTIVE AR TRAINING SYSTEMS

Table 5.5 gives the ranked extra requirements of an ARAT compared to a non-adaptive AR system. Again, the $\alpha=0,05$ significance threshold was not reached. Furthermore, the fact that the first three aspects have average ranks very close to each other displays that they are considered similarly important. Overall, it becomes visible that running an ARAT is considerably more complicated than running a non-adaptive AR system regarding the computations.

Furthermore, the increasing automation raises several new issues that human trainers can incorporate effortlessly, but might be an issue when setting up an ARAT like the timing of feedback. Lastly, the necessity for more sensors can be problematic as already current non-adaptive AR hardware systems are seen as too big and heavy (Haagsman, 2017), which would increase with more sensors.

A central result in this ranking is the need for a specific evaluation algorithm for every assembly step, which seven out of the nine respondents of the final round ranked this amongst the two most important requirements.

The most important requirement that did not get ranked is the necessity to define in which context the system has to work as having this too broadly defined would exceed the computational capabilities and having it too narrowly would lead to limited applicability (see Table A.6 in Appendix G).

Regarding this ranking, the interviewees did not have any further aspects. Also in this ranking, no evidence could be found in any listing supporting the relevance of ELT.

5.3 Identified future systems for the training phases (RQ)

The aforementioned results already offer valuable implications regarding the main RQ. Nevertheless, the Delphi participants were also asked directly which technology they prefer for which TWI step. The following sections will, therefore, describe the proposed systems and the assembly characteristics necessary for implementing ARATs.

5.3.1 Systems proposed

The experts expressed their opinion on which method is the best for each step of the TWI methodology (see section 2.1.2) As the majority of AR training systems are designed with the intention to reduce the necessity for human trainers (see section 2.4), the level of trainer inclusion the experts would put into their method choices has also been assessed. At first, these questions were asked without looking at the interconnections between those steps, so every step for itself. The results of this are illustrated in Figure 5.1. It becomes apparent that taken solely in three out of the four steps one solution can be

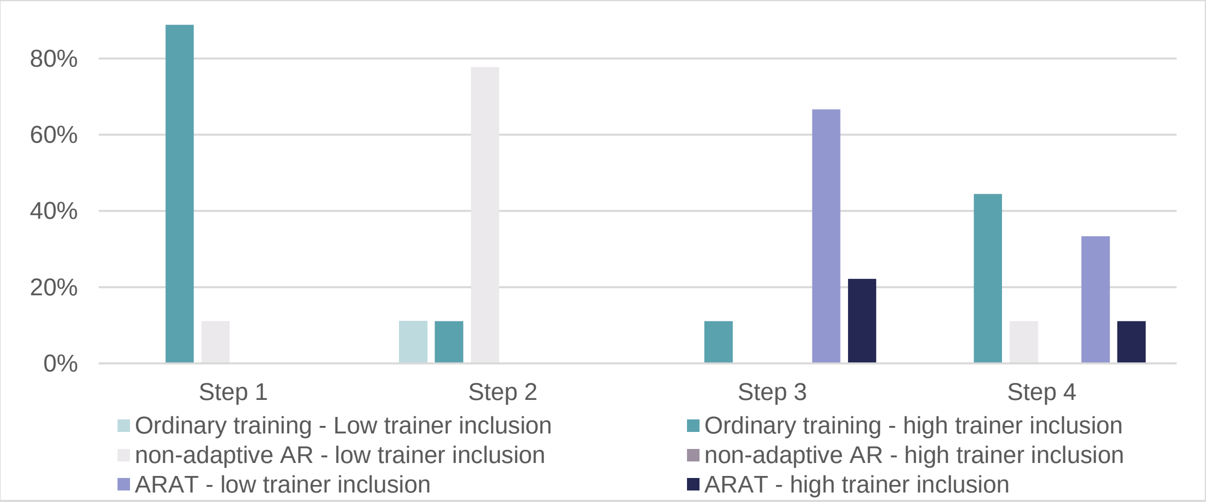

FIGURE 5.1: METHODOLOGY PREFERENCES WITHOUT INTERACTIONS

seen as the best choice. Only on the follow-up, the experts do not have a clear preference between the ordinary method and an ARAT. However, the different solutions for each consecutive step indicate a complicated design of interfaces to put those choices into a system.

Therefore, the same questions were also given to the experts with taking into account all possible interconnections and interactions between the different steps, so as a full TWI system. The results of this are visualized in Figure 5.2 and Table 5.6.

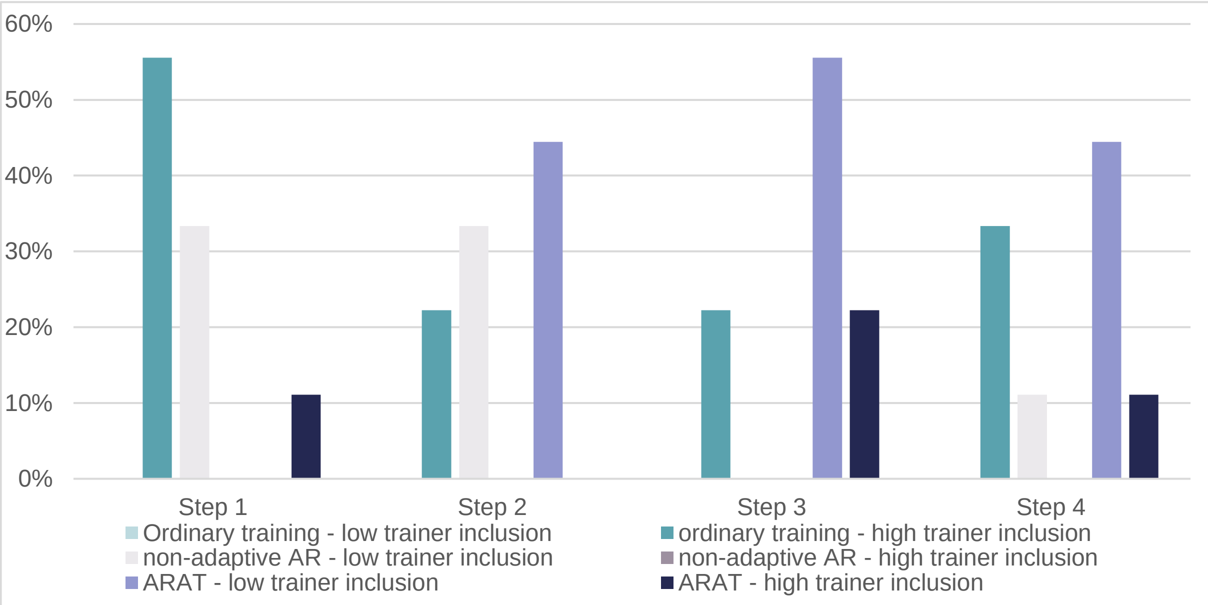

FIGURE 5.2: METHODOLOGY PREFERENCES WITH INTERACTIONS

Figure 5.2 displays the overall choices for each TWI step from a systemic perspective. Although preferences are still visible, they appear weaker for the systems. This shows that the system design is problematic. Regarding the preferences, the experts still see ordinary methods at least for the first step as the best method. Although for all other steps ARATs dominate, this supports the impression that a full substitution of ordinary training methods is improbable.

Table 5.6 illustrates the overall training systems that experts proposed and the share of experts who proposed them. The choices displayed show that like for the methodology itself especially the outer parts of the TWI methodology seem to be rather subjective choices (compare section 5.1.1). When only looking at the central steps 2 and 3, it becomes apparent that 44% of the experts prefer those to be both covered in ARATs and 33% speak for covering step 2 in non-adaptive AR and step 3 with an ARAT. This displays that although overall the number of suggested different systems is high, the judgement on the fundamental steps displays some unity.

The comparison of the two questions illustrates that there is no best practice established yet. While when taken solely, in most cases one method still had a clear majority, the answers for the system were diverse. However, it is still visible that overall the experts mainly see the strength of ARATs in

Method Choice	Steps				Steps				Steps				Steps				Steps				Steps				Steps			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Ordinary	x	x	x	x	x				x				x			x												
Non-adapt. AR										x		x		x			x				x	x						
ARAT						x	x	x			x				x			x	x	x			x	x	x	x	x	x
Share	22%				22%				11%				11%				11%				11%				11%			

TABLE 5.6: TRAINING SYSTEM CHOICES

step 3 (Try-out performance) of the TWI methodology, while the less interactive step 2 (present the operation) is still seen to be well representable in non-adaptive AR. When seen in a system however many experts already prefer ARATs for step 2, possibly to have a direct carry-over to step 3. Furthermore, it has to be noted that one participant appeared to have a high inconsistency between his answers in those two questions indicating a possible misunderstanding. If he would be excluded, the systemic results would be speaking less in favour of ordinary training while the results in Figure 5.1 would remain rather unchanged.

It becomes apparent that also the experts see modern technologies as a measure to reduce the dependence on the trainer in the training system. Interestingly, the only one of the systemic solutions that some experts combined with a still high trainer inclusion was the more interactive ARAT, where it could have been expected that a trainer would be needed the least. It is unclear how the high trainer inclusion was envisaged by the participants. However, opposing to what Ferrati *et al.* (2019) found out, in this investigation nobody believed that AR would completely replace human trainers.

5.3.2 Necessary assembly characteristics

As established in section 2.2.1, the assembly characteristics were expected to influence the methodological decision. Therefore, Table 5.7 gives the ranking of the two most important characteristics of an assembly supporting the implementation of an ARAT as ranked in the Delphi method. These characteristics give implications under which conditions in the assembly the strengths of ARATs come into play according to the experts. It becomes apparent that both characteristics are factors making an assembly more challenging.

Rank	Assembly characteristic
1	Fast reactions to errors are crucial
2	High product complexity

TABLE 5.7: NECESSARY CHARACTERISTICS OF ASSEMBLY PROCESSES FOR ARAT

As only two characteristics were recognised as important by the majority in the second round, Kendall's w could not be applied. The rankings were however very diverse.

The most important characteristics that have not been shortlisted were the possibilities to clearly define and closely follow a specific training process and to state which information should be checked when (see Table A.5 in Appendix G).

No given characteristic supports the relevance of ELT for ARATs.

Furthermore, the experts were confronted with the characteristics of two made-up case companies where they had to decide what kind of AR to recommend to them (see Appendix E). The main traits of these companies are summarized in Table 5.8. When being asked the first time, 60% recommended ARATs to case companies A. Interestingly, the information about the arguments others provided apparently changed their opinions as in the last round everyone recommended to use non-adaptive AR. For case company B, every expert recommended using ARATs in both rounds. The justifications for those decisions are given in Table A.7and Table A.8 in Appendix G.

Case Company A	High volume low-value low complexity product	looking for quick standardized and scalable training to prepare off the job for assembly line work
Case Company B	Low volume high-value high complexity product	looking for personalized training and long-term monitoring on the job with no mistake tolerance

TABLE 5.8: MAIN TRAITS EXEMPLARY CASE COMPANIES

6 Discussion

In the following section, the previously given results of the study and the limitations of the study will be discussed.

6.1 Review of the results

This study aimed at exploring which phases of industrial assembly training could be assisted best by using which kind of AR system. The main reason for the methodological selection was the hypothesis that there is a substantial gap between theory and practice of assembly training (see SQ2). This research suggests falsifying this by showing companies use own interpretations of methods close to what theory suggests them. As the companies investigated are all participating in the HAN umbrella project researching AR in assembly, they might be closer to research than many other comparable companies.

However, not all training theories are equally relevant in assembly practice. All examined companies followed the TWI theory. The ELT, in research also well renowned, seems to not practically not relevant in assembly training. This then opens a gap between theory and practice as Herbert *et al.* (2018) base their rationale for ARATs also on the ELT. As companies are very keen to get a perfect assembly done by their employees (interviewees 1.1, 1.2, 2.3), it might be a barrier to convince a company of a technology built on a theory they do not support. This issue has not yet been recognised neither in literature nor by a participant of this research.

However, ARATs do not enhance experiences of a level unknown at the companies, but rather imitate the experiences trainees are already getting nowadays during ordinary training in interaction with the trainer. Therefore, they are relying on a weak form of ELT that is already practised and used in companies, possibly unconsciously. Therefore, it can be assumed that companies would still acknowledge the potential ARATs offer, even though they do not completely agree with all implications of the ELT.

The hypothesized values of the assembly products in companies interested in implementing AR in assembly training turned out as expected for the case companies. Both case companies assembled products of a wide value range which mostly laid in the identified middle. This supports the theory formulated in section 2.2.1 of a relationship between product value and the virtuality of the training approach. However, it is not generalizable yet.

When looking at the Delphi results illustrated in Table 5.3 to Table 5.5, it becomes apparent that the consensus is statistically significant for only one ranking (see section 5.2). However, as this research is explorative and the Delphi method purposefully gathered experts from several different fields with different opinions, a statistical significance was improbable to be achieved. When clustered in the expert groups, the measures are higher, but no ranking is significant as the critical values are higher for fewer respondents (see Appendix H). The established ranks, therefore, should not be taken as clear and unmistakable lists, but rather as implication on the importance of factors. A factor being ranked first is not necessarily the most important but can be still seen as a very important factor to check.

The aim of this research was to provide companies thinking about implementing AR into their assembly training with the most important arguments for and against the different solutions. As the companies themselves will have to figure out which of the given aspects is more important for them, a statistically significant ranking is not necessary. Through the lists established by this research, companies get a good overview of the most important points that has been built upon the knowledge and experience of various experts in the relevant fields. Through the system answers some system blueprints are provided (see section 5.3.1). Therefore, companies now have a valid decision aid (see Appendix I) that can help them in a future decision on the implementation of AR in their assembly training. Furthermore, system developers get an impression on practical requirements and desires companies have for an AR assembly training system.

The established rankings also illustrate the problems and wishes that companies have in their assembly training. The list of identified advantages of the non-adaptive AR systems, therefore, seems to entail some characteristics that might not (yet) be realistically achievable (see Table 5.3). When looking at state-of-the-art systems like Aouam *et al.* (2018) or Werrlich *et al.* (2018), it becomes visible that also modern non-adaptive systems do not yet offer the degree of self-sustainability that some experts saw in them during the Delphi questionnaire. Similar phenomena were also noticed in other rankings. Therefore, the results might be biased to a certain extent.

6.2 Limitations

Like all pieces of research, this research has several limitations in its practical use.

First, the scope and time limits of a master thesis form a limitation. Although it was tried to create a big and well-balanced pool of experts and case companies, the cases still lack diversity so that the influence of different assembly characteristics on the optimal choice of technology could not be studied thoroughly and the results are not fully generalizable. The amount of Delphi rounds had to be limited to three, while it would have been helpful to send out especially the rankings more often like Schmidt *et al.* (2001) did in the hope to achieve stronger consensuses.

Furthermore, this research only investigates the possibilities of AR in assembly training. However, the AR technology reportedly also offers benefits in other occasions of an assembly like guidance, repair or maintenance (see Webel *et al.*, 2012; Haagsman, 2017). It can be assumed that also here ARATs would be having big advantages compared to non-adaptive AR systems and synergetic effects could be realisable. As a limitation on one aspect had to be taken to keep the research precise and in scope, those aspects could not be assessed in this research.

Another limitation lays in the explorative nature of the research. Although the topic of ARATs has been covered by other researchers, it still has not been realised yet in practice. Therefore even the best experts can only give hypotheses of what those systems will be capable of. As already described in the preceding section 6.1, this may lead to unrealistic daydreaming about features, but also has the direct consequence that straight away this research cannot be practically used or verified. However, it offers a showcase of possible capabilities of an ARAT and therefore can help a company to decide whether or not to invest into further research to design such a system and a developer what to focus on.

The used methodology in the Delphi investigation ensured that only aspects given by the experts were part of the process. This had the advantage of having first-hand aspects to be assessed without influence from the researcher, but also means that all arguments given have the same weight, independent of where it came from. This bears the risk that the participants without technology expertise envision technologically unrealistic capabilities of a system, which afterwards get voted to big importance as they would indeed solve major issues. Consequently, unrealistic aspects could get ranked. This could not have been prevented as a Delphi investigation intentionally aims at gathering unusual arguments in order to open new impressions and let the experts subsequently assess those. The unrealistic arguments then might have been voted high by some participants because of their value-added and voted low for being unachievable by others, but those intentions are not visible to the researcher. The dominance of technology experts in the panel was intended to moderate this. But as technology always got pushed by people pursuing what others claimed impossible, those yet unrealistic aspects might also enhance the creativity and motivate to find a solution making those possible.

Another methodological limitation lays in the design of the Delphi questionnaires. While the online questionnaire as data collection method offers convenience for the experts, it also entails some uncertainties for the researcher. As the answering of the questionnaire is impersonal, the researcher cannot be completely sure that the expert has filled it in in person. Furthermore, the written form of communication carries the risk of misunderstandings. While for the open question first-round questionnaire this could still be moderated through a validation of the groupings in the second round, it was impossible to ensure that the experts understood the closed ranking questions correctly, which might have interfered with their judgement.

Furthermore, the selection and narrowing down of the aspects in the second Delphi round was, on the one hand, a necessary part of the methodology to not have exceedingly long rankings, but meant on the other hand that several aspects got disregarded in the process. Especially regarding those close to the 50 % threshold, this means that potentially important aspects did not get ranked. However, in order to provide possibly interested readers with a full overview of the results, all answers given can be seen in Appendix G.

Lastly, a rather general limitation can be seen in the question of how much manual assembly will still be needed in the future. As Biao *et al.* (2011) emphasized, increasing labour costs are pushing the development further towards automating the former manual assembly operations. If this process would be continuing, the necessity for manual assembly and, therefore, also for workers trained in assembly will be decreasing. However, this does not mean that manual assembly will be completely disappearing as also the strong research interest in it implies. Lim and Hoffmann (2015) state that manual assembly is still crucial to production, especially where production numbers are rather small or much manipulation of the product is required. Therefore, the author does not think that manual assembly research is heading towards a dead end. Nevertheless, the developments in this regard will have to be followed.

7 Conclusion and further research

The aim of this research was to establish for which phases of industrial assembly training which kind of AR technology offers the most beneficial assistance. Through a multiple case study in companies currently working out AR assembly training systems, the current status quo of assembly training has been analysed. Afterwards, a Delphi investigation among 10 experts ranked the advantages and disadvantages of the different systems according to their importance and proposed options for a training system using AR technology.

The results achieved offer an opportunity for companies to figure out thoroughly whether or not to invest in implementing AR technologies in their assembly training and which kind of system is most beneficial for their circumstances.

Through the case study researching companies already working on the implementation of AR in assembly training, it was ensured that this research is covering problems and settings practically relevant in industrial assembly. The outcomes reveal that the TWI methodology plays a central role in a company's assembly training while the ELT as another highly relevant theory in research does not seem to be relevant in corporate practice (SQ1&2). Furthermore, it became apparent that mainly the lack of qualified personnel to train new employees and the pursuit to use them efficiently motivates companies to research technological solutions like AR in their assembly training.

The practical systems currently under development in the case companies proved to be non-adaptive mobile AR training systems using HMDs. However, both companies acknowledged the missing adaptability as problematic and another company is already developing an ARAT-like system for non-mobile workstations (SQ3).

Those key learnings from the case study were subsequently entered into a Delphi investigation with experts from AR, AI, assembly training and training methods. With the help of those experts, rankings of the most important advantages and disadvantages of non-adaptive AR in training (SQ4), of ARATs compared to non-adaptive AR training systems (SQ5) as well as the extra requirements of ARATs compared to non-adaptive AR have been established (SQ6). Those rankings can help companies interested in implementing AR into assembly training to weight up the system solutions for their assembly.

Furthermore, the experts also gave their preferred system choice for each step of the practically used TWI methodology. Through this, it became apparent that especially in a systemic perspective the experts see a lot of potential in ARATs for the future, but a human trainer should not get completely obsolete.

As it is dealing with technology not yet readily available, this research opens several interesting opportunities for future research.

One of the main aspects of this research was to gather the opinions from experts on what the technological advancement of ARATs can lead to in practical training. The result, lists of advantages those systems will bring with them can also be interpreted as a target list for developers to develop useful programmes. Therefore, with the growth of computer possibilities, research can then develop and test system solutions based on such lists of desired capabilities. Furthermore, it should also be tested which of the established desired system capabilities are practically possible.

Before it will be possible to practically research ARATs in a factory setting, future researchers can focus on the importance of specific assembly characteristics or settings. For example, it can be investigated whether an assembly training system for a worker on the assembly line differs from a system for an assembly worker in a manufacturing workshop. Haagsman (2017) provided a list of 17 characteristics influencing the general potential of using AR. This could be used as a basis to investigate the influence of those factors on the choice of AR technology.

This thesis focused on the characteristic of product value and its hypothesized influence on the technology choice. As the literature-based proposed relationship between the value and the virtuality of the training could not be fully verified with the only two case companies in this research, further research could more thoroughly investigate whether and to which extent this relationship can be observed in practice.

Another focus of this research has been laid on assembly training, so the worker's knowledge acquisition. However, as it is visible from the work of e.g. Haagsman (2017), Huenerfauth (2014) and Webel *et al.* (2012), AR offers a significant potential not only in assembly training but also in assembly guidance and maintenance. It can be assumed that sections of the investigated AR training systems could also add value in those areas, either in a synergistic symbiosis with AR training or taken solely. However, there is no research reviewing the possibilities of the different AR systems in those assembly contexts yet which opens up another area to be covered by future research.

This research also revealed some insights and ideas that originally have not been asked, but might nevertheless be starting points for future research. Interviewee 2.5 proposed to enhance the acceptance of ARAT systems by implementing a virtual trainer delivering the feedback. Future research might develop such a system and subsequently test whether a virtual person is increasing the acceptance and efficiency of the system's instructions.

Furthermore, interviewee 2.3 envisaged an AR training system entailing many main factors of the learning factory (see section 2.2.2) for his company with building up a dummy assembly line section for training purposes. Future research could develop and test such a combined system in a controlled environment to assess the synergies of combining AR and learning factories for both kinds of AR training systems, non-adaptive AR training systems and ARATs.

The direct effect of this research on industrial practice in assembly training will be limited as AR technology is still under development. However, the author hopes to have provided companies interested in modernizing their assembly training with a good decision aid (see Appendix I) on which circumstances support an investment in which kind of system and system developers with some insights from practice about what functionalities are desired in AR assembly training systems. Therefore, it is believed that the consequences and relevance of this research in practice will be visible in the future when companies underwent the development process.

For theory, this thesis gains its relevance by giving an overview of what kinds of AR systems have already been developed, what the outcomes of those were and how those could be further developed. Not only can this research make scientists aware of the opportunities non-adaptive AR and ARATs have in assembly training as motivation to intensify research, but also gives this research a blueprint of which kinds of systems are the best-suited for which environment. It can, therefore, help researchers to find the right setting to develop their systems in.

Overall, it became apparent that AR systems have a lot of potential in industrial assembly training, but a long development path still lays ahead. Through this research, another small step in this path has been taken. The future will reveal how big this step was.

8 List of references

Abele, E. (2016) 'Learning factory', *CIRP Encyclopedia of Production Engineering*, pp. 1-5.

Abele, E., Chryssolouris, G., Sihn, W., Metternich, J., ElMaraghy, H., Seliger, G., Sivard, G., ElMaraghy, W., Hummel, V. and Tisch, M. (2017) 'Learning factories for future oriented research and education in manufacturing', *CIRP annals*, 66(2), pp. 803-826.

Adler, M. and Ziglio, E. (1996) *Gazing into the oracle: The Delphi method and its application to social policy and public health*. Jessica Kingsley Publishers.

Ahmad, R., Masse, C., Jituri, S., Doucette, J. and Mertiny, P. (2018) 'Alberta Learning Factory for training reconfigurable assembly process value stream mapping', *Procedia Manufacturing*, 23, pp. 237-242.

Al-Ahmari, A., Ameen, W., Abidi, M.H. and Mian, S.H. (2018) 'Evaluation of 3D printing approach for manual assembly training', *International Journal of Industrial Ergonomics*, 66, pp. 57-62.

Aleven, V., McLaren, B.M. and Sewall, J. (2009) 'Scaling up programming by demonstration for intelligent tutoring systems development: An open-access web site for middle school mathematics learning', *IEEE Transactions on Learning Technologies*, 2(2), pp. 64-78.

Alqahtani, F. and Ramzan, N. (2019) 'Comparison and Efficacy of Synergistic Intelligent Tutoring Systems with Human Physiological Response', *Sensors*, 19(3), p. 460.

Anastassova, M. and Burkhardt, J.-M. (2009) 'Automotive technicians' training as a community-of-practice: Implications for the design of an augmented reality teaching aid', *Applied ergonomics*, 40(4), pp. 713-721.

Aouam, D., Benbelkacem, S., Zenati, N., Zakaria, S. and Meftah, Z. (2018) *2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS)*. 24-25 Oct. 2018.

Azuma, R.T. (1997) 'A survey of augmented reality', *Presence: Teleoperators & Virtual Environments*, 6(4), pp. 355-385.

Biao, Z., Jianjun, W., Rossano, G., Martinez, C. and Kock, S. (2011) 'Vision-guided robot alignment for scalable, flexible assembly automation'. pp. 944-951.

Billinghurst, M. and Dünser, A. (2012) 'Augmented Reality in the Classroom', *Computer*, 45(7), pp. 56-63.

BMW Group (2019) 'Absolutely real: Virtual and augmented reality open new avenues in the BMW Group Production System'. 09/04/2019. press.bmwgroup.com. Available at: https://www.press.bmwgroup.com/global/article/detail/T0294345EN/absolutely-real:-virtual-and-augmented-reality-open-new-avenues-in-the-bmw-group-production-system?language=en.

Boud, A.C., Haniff, D.J., Baber, C. and Steiner, S.J. (1999) *1999 IEEE International Conference on Information Visualization (Cat. No. PR00210)*. 14-16 July 1999.

Bowen, F., Cousins, P., Lamming, R. and Faruk, A. (2006) 'Horses for courses: explaining the gap between the theory and practice of green supply', in *Greening the supply chain*. Springer, pp. 151-172.

Bower, G.G. (2014) 'Theory and practice: Utilizing Dewey's experiential learning theory to implement a 5k road race', *Journal of Hospitality, Leisure, Sport & Tourism Education*, 15(1), pp. 61-67.

Brauer, M. (2014) *An der Hochschule lehren: Praktische Ratschläge, Tricks und Lehrmethoden*. Springer-Verlag.

Brusilovsky, P., Schwarz, E. and Weber, G. (1996) *International conference on intelligent tutoring systems*. Springer.

Bryman, A. and Bell, E. (2007) *Business research methods*. 2nd ed.. edn. Oxford: Oxford : Oxford University Press.

Cachay, J., Wennemer, J., Abele, E. and Tenberg, R. (2012) 'Study on Action-Oriented Learning with a Learning Factory Approach', *Procedia - Social and Behavioral Sciences*, 55, pp. 1144-1153.

Caudell, T.P. and Mizell, D.W. (1992) *Proceedings of the twenty-fifth Hawaii international conference on system sciences*. IEEE.

Charoenseang, S. and Panjan, S. 6773 (2011) '5-Finger exoskeleton for assembly training in augmented reality'. pp. 30-39.

Coad, A., Pellegrino, G. and Savona, M. (2016) 'Barriers to innovation and firm productivity', *Economics of Innovation and New Technology*, 25(3), pp. 321-334.

Cohen, L., Duboé, P., Buvat, J., Meltont, D., Khadikar, A. and Shah, H. (2018) *Augmented and Virtual Reality in Operations*. Capgemini Research Institute. [Online]. Available at: https://www.capgemini.com/de-de/resources/augmented-reality-virtual-reality-unternehmen-studie/#.

Corbin, J.M. and Strauss, A. (1990) 'Grounded theory research: Procedures, canons, and evaluative criteria', *Qualitative sociology*, 13(1), pp. 3-21.

Danielsson, O., Syberfeldt, A., Brewster, R. and Wang, L. 63 (2017) 'Assessing Instructions in Augmented Reality for Human-robot Collaborative Assembly by Using Demonstrators' *Procedia CIRP*. pp. 89-94.

Daponte, P., De Vito, L., Picariello, F. and Riccio, M. (2014) 'State of the art and future developments of the Augmented Reality for measurement applications', *Measurement*, 57, pp. 53-70.

De Vin, L.J. and Jacobsson, L. (2017) 'Karlstad lean factory: an instructional factory for game-based lean manufacturing training', *Production & Manufacturing Research*, 5(1), pp. 268-283.

Dermeval, D., Paiva, R., Bittencourt, I.I., Vassileva, J. and Borges, D. (2018) 'Authoring Tools for Designing Intelligent Tutoring Systems: a Systematic Review of the Literature', *International Journal of Artificial Intelligence in Education*, 28(3), pp. 336-384.

Dernova, M. (2015) 'Experiential Learning Theory As One Of The Foundations Of Adult Learning Practice Worldwide', *Comparative Professional Pedagogy*, 5(2), pp. 52-57.

Dewey, J. (1938) *Experience and education*. New York: Collier Books.

Dinero, D. (2005) *Training Within Industry: The Foundation of Lean*. Taylor & Francis.

Dooley, C.R. (1945) 'The Training within Industry Report 1940-1945', *Advances in Developing Human Resources*, 3(2), pp. 127-289.

Dünser, A., Steinbügl, K., Kaufmann, H. and Glück, J. (2006) 'Virtual and augmented reality as spatial ability training tools', *Proceedings of the 7th ACM SIGCHI New Zealand chapter's international conference on Computer-human interaction: design centered HCI*. Christchurch, New Zealand. 1152776: ACM, pp. 125-132.

Evens, M. and Michael, J. (2006) *One-on-one tutoring by humans and computers*. Psychology Press.

Falck, A.-C., Tarrar, M., Mattsson, S., Andersson, L., Rosenqvist, M. and Söderberg, R. (2017) 'Assessment of manual assembly complexity: a theoretical and empirical comparison of two methods', *International Journal of Production Research*, 55(24), pp. 7237-7250.

Ferrati, F., Erkoyuncu, J.A. and Court, S. (2019) 'Developing an Augmented Reality Based Training Demonstrator for Manufacturing Cherry Pickers', *Procedia CIRP*, 81, pp. 803-808.

Gavish, N., Gutiérrez, T., Webel, S., Rodríguez, J., Peveri, M., Bockholt, U. and Tecchia, F. (2015) 'Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks', *Interactive Learning Environments*, 23(6), pp. 778-798.

Gilbert, S.B., Blessing, S.B. and Guo, E. (2015) 'Authoring effective embedded tutors: An overview of the extensible problem specific tutor (xPST) system', *International Journal of Artificial Intelligence in Education*, 25(3), pp. 428-454.

Grigoriadou, M., Papanikolaou, K., Kornilakis, H. and Magoulas, G. (2001) *Proceedings of 3rd Workshop on Adaptive Hypertext and Hypermedia*.

Günnel, T. (2019) *Virtual und Augmented Reality: Wofür BMW die Technik einsetzt*. Available at: https://www.automobil-industrie.vogel.de/virtual-und-augmented-reality-wofuer-bmw-die-technik-einsetzt-a-780364/ (Accessed: 29/05).

Haagsman, L. (2017) *The Power of Information: Exploring the influence of assembly context on the potential of augmented reality in assembly operations*. University of Groningen.

Hair, J.F. (2007) *Research methods for business*. Chichester, England

Hoboken, N.J. : John Wiley & Sons Ltd.

Hammerstingl, V. and Reinhart, G. (2017) 'Fähigkeiten in der Montage'.

Hasson, F., Keeney, S. and McKenna, H. (2000) 'Research guidelines for the Delphi survey technique', *Journal of advanced nursing*, 32(4), pp. 1008-1015.

Herbert, B., Ens, B., Weerasinghe, A., Billinghurst, M. and Wigley, G. (2018) 'Design considerations for combining augmented reality with intelligent tutors', *Computers & Graphics*, 77, pp. 166-182.

Ho, N., Wong, P.-M., Chua, M. and Chui, C.-K. (2018) 'Virtual reality training for assembly of hybrid medical devices', *Multimedia Tools and Applications*, 77(23), pp. 30651-30682.

Hořejší, P. (2015) 'Augmented Reality System for Virtual Training of Parts Assembly', *Procedia Engineering*, 100, pp. 699-706.

Hou, L. and Wang, X. (2012) 'A Study on the Benefits of Augmented Reality in Retaining Working Memory in Assembly Tasks: A Focus on Differences in Gender', *Automation in Construction*.

Hou, L., Wang, X., Bernold, L. and Love, P.E.D. (2013) 'Using animated augmented reality to cognitively guide assembly', *Journal of Computing in Civil Engineering*, 27(5), pp. 439-451.

Huenerfauth, A.M. (2014) 'Mobile Technology Applications for Manufacturing, Reduction of Muda (Waste) and the Effect on Manufacturing Economy and Efficiency', *International Journal of Interactive Mobile Technologies*, 8(4), pp. 20-23.

Ivanova, M. (2013) *2013 International Conference on Interactive Collaborative Learning (ICL)*. IEEE.

Jee, H.-K., Lim, S., Youn, J. and Lee, J. (2014) 'An augmented reality-based authoring tool for E-learning applications', *Multimedia Tools and Applications*, 68(2), pp. 225-235.

Jee, H., Lim, S., Youn, J. and Lee, J. (2011) *2011 International Conference on Information Science and Applications*. 26-29 April 2011.

Jetter, J., Eimecke, J. and Rese, A. (2018) 'Augmented reality tools for industrial applications: What are potential key performance indicators and who benefits?', *Computers in Human Behavior*, 87, pp. 18-33.

Jueptner, M., Frith, C., Brooks, D., Frackowiak, R. and Passingham, R. (1997) 'Anatomy of motor learning. II. Subcortical structures and learning by trial and error', *Journal of neurophysiology*, 77(3), pp. 1325-1337.

Karlsson, C. (2016) *Research methods for operations management*. Second edition.. edn. Oxon : Routledge.

Kaufmann, H. (2004) *Geometry Education with Augmented Reality*. Technische Universität Wien.

Kokku, R., Sundararajan, S., Dey, P., Sindhgatta, R., Nitta, S. and Sengupta, B. 2018- (2018) 'Augmenting Classrooms with AI for Personalized Education'. pp. 6976-6980.

Kolb, D.A. (1984) *Experiential learning : experience as the source of learning and development*. Englewood Cliffs, N.J.: Prentice-Hall.

Kolb, D.A., Boyatzis, R.E. and Mainemelis, C. (2001) 'Experiential learning theory: Previous research and new directions', *Perspectives on thinking, learning, and cognitive styles*, 1(8), pp. 227-247.

Kreft, S., Gausemeier, J. and Matysczok, C. 2 (2009) 'Towards Wearable Augmented Reality in automotive assembly training'. pp. 1537-1547.

Kulik, J.A. and Fletcher, J.D. (2016) 'Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review', *Review of Educational Research*, 86(1), pp. 42-78.

Laick, S. (2012) 'Using Delphi methodology in information system research', *International Journal of Management Cases*, 14(4), pp. 261-268.

Langley, A., Lawson, G., Hermawati, S., D'Cruz, M., Apold, J., Arlt, F. and Mura, K. (2016) 'Establishing the usability of a virtual training system for assembly operations within the automotive industry', *Human Factors and Ergonomics in Manufacturing & Service Industries*, 26(6), pp. 667-679.

Leblanc, F., Champagne, B.J., Augestad, K.M., Neary, P.C., Senagore, A.J., Ellis, C.N. and Delaney, C.P. (2010) 'A Comparison of Human Cadaver and Augmented Reality Simulator Models for Straight Laparoscopic Colorectal Skills Acquisition Training', *Journal of the American College of Surgeons*, 211(2), pp. 250-255.

Lee, K. (2012) 'Augmented Reality in Education and Training', *TechTrends*, 56(2), pp. 13-21.

Lewin, K. (1951) *Field theory in social science : selected theoretical papers*. London: London : Tavistock Publications.

Li, S., Peng, T., Xu, C., Fu, Y. and Liu, Y. 5622 (2009) 'A mixed reality-based assembly verification and training platform'. pp. 576-585.

Lim, J.T.A. and Hoffmann, E.R. (2015) 'Strategies in performing a manual assembly task', *International Journal of Industrial Ergonomics*, 50, pp. 121-129.

Linstone, H.A. and Turoff, M. (1975) *The Delphi method : techniques and applications*. Reading, Mass.: Reading, Mass. : Addison-Wesley Pub. Co., Advanced Book Program.

Liu, T.Y. (2009) 'A context-aware ubiquitous learning environment for language listening and speaking', *Journal of Computer Assisted Learning*, 25(6), pp. 515-527.

Liu, Y., Li, S. and Wang, J. (2015a) 'Assembly auxiliary system for narrow cabins of spacecraft', *Chinese Journal Of Mechanical Engineering*, 28(5), pp. 1080-1088.

Liu, Y., Li, S., Wang, J., Zeng, H. and Lu, J. (2015b) 'A computer vision-based assistant system for the assembly of narrow cabin products', *The International Journal of Advanced Manufacturing Technology*, 76(1), pp. 281-293.

Lucrecia, M., Cecilia, S., Patricia, P. and Sandra, B. (2013) *2013 International Conference on Collaboration Technologies and Systems (CTS)*. 20-24 May 2013.

Martín-Gutiérrez, J., Luís Saorín, J., Contero, M., Alcañiz, M., Pérez-López, D.C. and Ortega, M. (2010) 'Design and validation of an augmented book for spatial abilities development in engineering students', *Computers & Graphics*, 34(1), pp. 77-91.

Meredith, J. (1998) 'Building operations management theory through case and field research', *Journal of Operations Management*, 16(4), pp. 441-454.

Milgram, P., Kishino, F.J.I.T.o.I. and Systems (1994) 'A taxonomy of mixed reality visual displays', 77(12), pp. 1321-1329.

Milgram, P., Takemura, H., Utsumi, A. and Kishino, F. (1995) 'Augmented Reality: A class of displays on the reality-virtuality continuum'.

Mitrovic, A., Martin, B., Suraweera, P., Zakharov, K., Milik, N., Holland, J. and McGuigan, N. (2009) 'ASPIRE: an authoring system and deployment environment for constraint-based tutors', *International Journal of Artificial Intelligence in Education*, 19(2), pp. 155-188.

Morkos, B., Taiber, J., Summers, J., Mears, L., Fadel, G. and Rilka, T. (2012) 'Mobile devices within manufacturing environments: a BMW applicability study', *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 6(2), pp. 101-111.

Mura, M.D., Dini, G. and Failli, F. (2016) 'An Integrated Environment Based on Augmented Reality and Sensing Device for Manual Assembly Workstations', *Procedia CIRP*, 41, pp. 340-345.

Myers, M.D. (2009) *Qualitative research in business and management*. Los Angeles: SAGE.

Nash, B., Walker, A. and Chambers, T. (2018) 'A simulator based on virtual reality to dismantle a research reactor assembly using master-slave manipulators', *Annals of Nuclear Energy*, 120, pp. 1-7.

Nee, A.Y.C., Ong, S.K., Chryssolouris, G. and Mourtzis, D. (2012) 'Augmented reality applications in design and manufacturing', *CIRP Annals*, 61(2), pp. 657-679.

Neges, M., Adwernat, S. and Abramovici, M. (2018) 'Augmented Virtuality for maintenance training simulation under various stress conditions', *Procedia Manufacturing*, 19, pp. 171-178.

Nof, S.Y., Wilhelm, W.E. and Warnecke, H. (2012) *Industrial assembly*. Springer Science & Business Media.

Nöhring, F., Rieger, M., Erohin, O., Deuse, J. and Kuhlenkötter, B. 32 (2015) 'An Interdisciplinary and Hands-on Learning Approach for Industrial Assembly Systems'. pp. 109-114.

Novak-Marcincin, J., Barna, J., Janak, M. and Novakova-Marcincinova, L. (2013) 'Augmented Reality Aided Manufacturing', *Procedia Computer Science*, 25, pp. 23-31.

Okoli, C. and Pawlowski, S.D. (2004) 'The Delphi method as a research tool: an example, design considerations and applications', *Information & Management*, 42(1), pp. 15-29.

Ong, S.K., Yuan, M.L. and Nee, A.Y.C. (2008) 'Augmented reality applications in manufacturing: a survey', *International Journal of Production Research*, 46(10), pp. 2707-2742.

Petrovica, S. and Ekenel, H.K. (2016) *BIR Workshops*.

Pozzi, R., Noè, C. and Rossi, T. (2014) 'Experimenting ‘learn by doing’ and ‘learn by failing’', *European Journal of Engineering Education*, 40(1), pp. 1-13.

Ramteke, S.K. (2019) 'Innovation Strategies for a Global Manufacturing Business'.

Ranz, F., Komenda, T., Reisinger, G., Hold, P., Hummel, V. and Sihn, W. (2018) 'A Morphology of Human Robot Collaboration Systems for Industrial Assembly', *Procedia CIRP*, 72, pp. 99-104.

Rentzos, L., Papanastasiou, S., Papakostas, N. and Chryssolouris, G. (2013) 'Augmented Reality for Human-based Assembly: Using Product and Process Semantics', *IFAC Proceedings Volumes*, 46(15), pp. 98-101.

Santos, M.E.C., Chen, A., Taketomi, T., Yamamoto, G., Miyazaki, J. and Kato, H. (2014) 'Augmented Reality Learning Experiences: Survey of Prototype Design and Evaluation', *IEEE Transactions on Learning Technologies*, 7(1), pp. 38-56.

Sato, T. and Laughlin, D.D. (2018) 'Integrating Kolb's Experiential Learning Theory into a sport psychology classroom using a golf-putting activity', *Journal of Sport Psychology in Action*, 9(1), pp. 51-62.

Saunders, M.N.K., Lewis, P. and Thornhill, A. (2019) *Research methods for business students*. Eigth edition.. edn. Harlow: Harlow : Pearson.

Schmidt, R., Lyytinen, K., Keil, M. and Cule, P. (2001) 'Identifying software project risks: An international Delphi study', *Journal of management information systems*, 17(4), pp. 5-36.

Schmidt, R.C. (1997) 'Managing Delphi surveys using nonparametric statistical techniques', *decision Sciences*, 28(3), pp. 763-774.

Sekaran, U. and Bougie, R. (2016) *Research methods for business : a skill-building approach*. Seventh edition.. edn. Chichester, West Sussex, United Kingdom : John Wiley & Sons.

Serino, M., Cordrey, K., McLaughlin, L. and Milanaik, R.L. (2016) 'Pokémon Go and augmented virtual reality games: a cautionary commentary for parents and pediatricians', *Current opinion in pediatrics*, 28(5), pp. 673-677.

Siegel, S. (1988) *Nonparametric statistics for the behavioral sciences*. 2nd ed.. edn. New York: New York : McGraw-Hill.

Sleeman, D. and Brown, J.S. (1982) *Intelligent tutoring systems*. London: Academic Press.

Stricker, D., Klinker, G. and Reiners, D. (1999) *Proceedings of the international workshop on Augmented reality: placing artificial objects in real scenes: placing artificial objects in real scenes*. AK Peters, Ltd.

Syberfeldt, A., Holm, M., Danielsson, O., Wang, L. and Brewster, R.L. (2016) 'Support Systems on the Industrial Shop-floors of the Future – Operators' Perspective on Augmented Reality', *Procedia CIRP*, 44, pp. 108-113.

Tatić, D. and Tešić, B. (2017) 'The application of augmented reality technologies for the improvement of occupational safety in an industrial environment', *Computers in Industry*, 85, pp. 1-10.

Ternier, S., Klemke, R., Kalz, M., Van Ulzen, P. and Specht, M.J.J.U. (2012) 'ARLearn: Augmented Reality Meets Augmented Virtuality', 18(15), pp. 2143-2164.

VanLehn, K., Graesser, A.C., Jackson, G.T., Jordan, P., Olney, A. and Rosé, C.P. (2007) 'When are tutorial dialogues more effective than reading?', *Cognitive science*, 31(1), pp. 3-62.

Voss, C., Johnson, M. and Godsell, J. (2016) 'Case Research', in Karlsson, C. (ed.) *Research Methods for Operations Management*. 2nd edn., pp. 165-197.

Wang, X., Ong, S.K. and Nee, A.Y. (2016) 'A comprehensive survey of augmented reality assembly research', *Advances in Manufacturing*, 4(1), pp. 1-22.

Webel, S., Bockholt, U., Engelke, T., Gavish, N., Olbrich, M. and Preusche, C. (2012) 'An augmented reality training platform for assembly and maintenance skills', *Robotics and Autonomous Systems*.

Weber, G. and Brusilovsky, P. (2001) 'ELM-ART: An adaptive versatile system for Web-based instruction', *International Journal of Artificial Intelligence in Education (IJAIED)*, 12, pp. 351-384.

Werrlich, S., Lorber, C., Nguyen, P.A., Yanez, C.E.F. and Notni, G. 10909 (2018) 'Assembly training: Comparing the effects of head-mounted displays and face-to-face training'. pp. 462-476.

Westerfield, G., Mitrovic, A. and Billinghurst, M. (2015) 'Intelligent Augmented Reality Training for Motherboard Assembly', *International Journal of Artificial Intelligence in Education*, 25(1), pp. 157-172.

Xia, P., Lopes, A., Restivo, M. and Yao, Y. (2012) 'A new type haptics-based virtual environment system for assembly training of complex products', *The International Journal of Advanced Manufacturing Technology*, 58(1), pp. 379-396.

Xu, C., Li, S., Wang, J., Peng, T. and Xie, M. (2008) *Occlusion Handling in Augmented Reality System for Human-Assisted Assembly Task*. Berlin, Heidelberg: Berlin, Heidelberg: Springer Berlin Heidelberg.

Yamamoto, H. (1999) *IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No.99CH37028)*. 12-15 Oct. 1999.

Yin, R.K.-z. (2014) *Case study research : design and methods*. Fifth edition, [second printing]. edn. Thousand Oaks, CA: Sage Publications.

Appendices

Appendix A: Summary of literature on AR applications in assembly training

The following table summarizes chronologically ordered the different applications of AR assembly training systems in literature.

Publication	Approach	AR Hardware	AR Software	Purpose	research products	outcome
Boud et al. (1999)	Investigation of VR and AR potential for assembly training	HMD and headset connected to PC	Not specified	Compare the technologies and present the relative advantages of both as training tool	Water pump	AR assembly training significantly faster than conventional and VR training
Li et al. (2009)	Development of a MR assembly platform for planning, verification and training	VR glasses and camera system	Self-developed, adopted from Stricker et al. (1999) in C++	Superimposing virtual models, geometric features and information into a real assembly scene	Not specified	Platform could improve product quality, shorten development time and reduce costs (no evidence given)
Anastassova and Burkhardt (2009)	Qualitative evaluation of AR possibilities to assist in automotive service technicians training	Handheld device	Not specified	Exploration of how future maintenance staff training could be designed	Car to be main-tained	AR offers valuable possibilities to improve efficiency and effectiveness of the training
Kreft et al. (2009)	Framework for the development of a wearable AR system (combining wearable computing and AR)	HMD (by Carl Zeiss), wearable computer (OQO), data glove, sensors at product and Bluetooth headset	Self-developed, based on the wearable compu-tation application of “wearIT@ work”	More efficient training and guidance of mobile workers	Skoda car assembly training	Process model to systematically develop and evaluate AR enhancements

Charoenseang and Panjan (2011)	Virtual assembly training with force feedback simulating the product	Hand exo-skeleton, VR glasses, video camera, force sensors, markers	Graphics manager and vision manager	Assessing simulation possibilities for training without physical objects	Virtual peg hole puzzle	System possibility to perform virtual assembly training
Webel et al. (2012)	Multimodal AR-based training of maintenance and assembly skills	Tablet PCs, haptic bracelet	Self-developed, giving adaptive visual aids (AVAs)	Assessing the advantages of AR training systems (in maintenance)	Electro-mecha-nical actuator	AR trained people needed less aids, made less errors
Morkos et al. (2012)	Applicability study of mobile technology in BMW production	HMD, Tracking system, mobile computer	Only hypothe-tical	Asses the advantages and disadvantages in manufac-turing	Different BMW parts	Provide standardized training, reduce trainer costs
Huenerfauth (2014)	Reduction of Muda (waste) with mobile technology applications	None (just evaluates existing models)	None (just evaluates existing models)	Evaluation of mobile applications to reduce muda in manufacturing	None (just eva-luates existing models)	AR training mainly reduces unnecessary / excess motion
Liu et al. (2015a) & Liu et al. (2015b)	Mixed reality to plan, train, guide and statues inspect assembly and assess quality (assembly assistant system)	Computer system with camera and display	Self-developed in C++, based on improved ant colony algorithm	Holistic optimization of the assembly of narrow spacecraft cabins	Space-craft cabins	Proposed systems feasible & practical for cabin assembly
Hořejší (2015)	Virtual training using AR	Installed PC and camera on the ceiling	Self-developed based on “Unifeye”	Comparison of the effectiveness of AR training compared to paper-based training	Gully trap	First assembly quicker and learning of the assembly more efficient
Gavish et al. (2015)	AR and VR platforms in industrial maintenance and assembly	Tablet PC, haptic bracelet	Self-developed	Evaluation of their potential uses in the industrial environment	Part of an electronic actuator of a moto-rized modula-ting valve	AR and VR training groups required more training time, fewer unsolved errors in AR

Westerfield et al. (2015)	Combination of AR and intelligent tutoring systems (ITS)	HMD (Wrap 920AR), PC and ARToolkit fiducial markers	Self-developed based on the ASPIRE ITS software and 3D Studio Max	Comparison of different levels of AR assembly training	Mother-boards	Intelligent AR systems had 25% higher test scores on quality and were 30% faster in task performance
Werrlich et al. (2018)	Augmented Reality for engine assembly training	HMD (Microsoft HoloLens)	Self-developed with game engine “Unity 3D”	Comparison between AR and face-to-face training with real-life engine assembly	Three different internal combus-tion engines	10% less picking errors, 5% less assembly errors and 60% less rework needed, but 60% slower learning
BMW Group (2019)	Implemen-tation of AR in assembly training	HMD	Self-developed	Train employees more efficiently in lean principles	Engine assembly	Number of trainees per trainer has tripled
Ferrati *et al.* (2019)	AR in cherry picker assembly	HMD (Microsoft HoloLens)	Self-developed combina-tion of Unity and Vuforia	Enhance the training process, improve learning time and error rate	Hydraulic hoses assembly to valves	22% faster assembly; time to pick reduced by 26%; number of help requests and error rate reduced, better questionnaire results

TABLE A.1: COMPARISON AR ASSEMBLY TRAINING SYSTEMS

Appendix B: Case study protocol

Overview of the Case Study	Mission and goals	Background	In literature, there are several approaches on how to best train employees in new tasks. These approaches such as the experiential learning theory (ELT) (originally Dewey, 1938) or Training within industries (TWI) (Dooley, 1945) are proven to be beneficial and frequently followed in practice. However, it is assumed that trainers diverge often from those theories based on their own experiences. Therefore, the actual training practice should be observed before the Delphi method will evaluate the AR applicability for the phases.
		Mission of the study	Establishing how assembly training is performed in practice. The focus is laid here on the interaction between trainer and trainee and the feedback and adaptions of training approach made based on the trainee's performance.
		Goals to be achieved	Establish a realistic picture how staff are trained in practice and compare it to the underlying theory. This practice is later on compared to the AR system capabilities
	Why Case Study?	• Contemporary events and practices shall be investigated and case study research allows direct, real time phenomena to be studied (Karlsson, 2016) • Case research is widely used in operations management and augmented reality research (e.g. Anastassova and Burkhardt, 2009; Liu *et al.*, 2015b; Westerfield *et al.*, 2015)	
	Case Study hard facts	Questions	RQ: Which phases in the industrial assembly training could be assisted best by which kind of AR training system? → empirical case study questions: • Which (if any) training theory is followed? • Where (if at all) does practice deviate from the theory? Which are the motivations for the trainers to do so? • Why AR (and not VR/robots/…)? • How much interaction and feedback on performance is involved in the training? Which kinds of interaction are used? • How thoroughly are the trainees monitored during their assembly trials? • How much are the approaches adapted to the trainee's capabilities? • By how much do the attributes of the assembly product (esp. value) matter in the training approach and the benefits of adapting AR assistance?

		Hypotheses	• Theories are followed, but practical deviations are taken • Interaction between participants and the adaptation to the trainee's performance play an important role and are frequently used to a big extent • Trainees are closely monitored to assess their capabilities and adapt the training to them
		Propositions	• Although there are a number of theories, they always have to be at least partly adapted for practical usage • Training needs interaction and individual adaptations to be effective
	Theoretical framework for the case study, key readings	Amongst many theories, ELT and TWI proved to be two of the most influential for industrial staff training, including assembly training. Their influence on practice however has only rarely been assessed (Bower, 2014; Dernova, 2015; Sato and Laughlin, 2018) and yet never took reference to industrial training settings. As it is known that corporate practice often differs from theory (e.g. Bowen *et al.*, 2006), the actual practice has to be assessed. (see further section 2) **Key readings:** Dewey (1938), Kolb (1984), Dooley (1945), Dinero (2005)	
	Role of the protocol in guiding the researcher	The protocol aims at guiding the researcher through the case study at hand by ensuring that the intended outcomes can be achieved. It serves as a reminder what data and information are needed, how they have to be collected and recorded and how the findings have to be presented in the report. Furthermore it serves as research documentation to ensure reliability.	
Case setting	Unit of Analysis	Regular staff training for assembly of different products (without AR)	
	Choice philosophy	Pragmatic choice due to involvement of companies participating in research project by the HAN	
	Case criteria	Assembly	Companies should be assembling some sort of product
		Location	Companies should be located in the Netherlands to avoid excessive travelling
		Case variety	The assembled products should differ in their assembly characteristics
		AR interest	Companies should be interested in implementing AR systems in their assembly training or already did so
Case Selection	Unit of analysis	Training activities of the companies for the usage of the trained staff in assembly actions	
	Focus	Obtain contemporary data of how the training theory is performed in practice	
	No. of cases	two	
	Replication logic	Theoretical replication logic (aiming for a diversity in the cases)	
	Selection criteria (following	Boundaries	Companies interested in implementing AR in assembly staff training
		Sample frame	Industrial assembly companies in the Netherlands

	Yin, 2014; Karlsson, 2016)	Sample control	SMEs in competitive markets
		Replication logic	Theoretical replication, based on different assembly products values
Data collection procedures	Names of Contact persons for doing fieldwork	Case company 1	confidential
		Case company 2	confidential
	Data collection plan	Data sources	Company documents and protocols, corporate training practice, employee's impressions and opinions
		types of evidence	• Company Documentations • direct observations • shorter case study interviews(Yin, 2014)
		time planning	depending on the availability of the company representatives
	Expected preparation prior to fieldwork	Organizing	• Making appointment with company • organizing travels to company • informing participants for informed consent
		researcher preparation	• Preparation of protocol • gathering of questions to be answered • preparing questions for interviews • defining which question should be handled by which sources
		Physical preparation	• Charging and preparing recording devices • Prepare safe and encrypted storage of all (especially personal) data
	Planned Interaction with participants	Before data collection	• Statement of gratefulness and introduction of the researcher and his studies at NCL and RUG • Explanation of study background and purpose → Why interesting for managers? → Why this company? • Discuss consent form, explanation of participant's rights, introduction of recording device and purpose • Planned structure and time plan of the study • Explanation of data collection
		During data collection	• Researcher will passively observe and record training sessions and take notes • Interviews will contain formal personal verbal interaction which will be recorded
		After data collection	• Explanation of data processing and next steps • Preview of when and where the results will be available • Contact exchange for further questions or changes in preferences/consent and transcript check • Asking for comments and feedback on the researcher • Thank for participation

	Procedures for protecting human subjects	• Anonymization of all participants and companies in the report • creating informed consent for all participants • leaving all participants the possibility to withdraw their consent without any specific reasons • respecting all data protection laws

Data collection questions	Categorized following the five levels of questions (Yin, 2014)	• considering SQ1:
		  ○ Which (if any) training theory is followed? (**Level 2**)
		  ○ How much interaction and feedback on performance is involved in the training? (**Level 2**)
		  ○ How thoroughly are the trainees monitored during their assembly trials? (**Level 2**)
		  ○ How much are the approaches adapted to the trainee's capabilities? (**Level 2**)
		• considering SQ2:
		  ○ Where (if at all) does practice deviate from the theory? (**Level 2**)
		  ○ Which are the motivations for the trainers to do so? (**Level 3**)
		  ○ How much are the approaches adapted to the trainee's capabilities? (**Level 2**)
		  ○ Why AR and not another technology (e.g. VR, robots)? (**Level 2**)
		• Considering SQ3:
		  ○ What does the AR training system you are currently developing look like? (**Level 2**)
		• considering SQ4:
		  ○ Which advantages and disadvantages do you see in AR compared to your current assembly training? (**Level 4**)
		• considering SQ5:
		  ○ Which advantages and disadvantages do you see in ARATs compared to non-adaptive AR training systems (**Level 4**)
		• considering SQ6:
		  ○ Which additional requirements do you see for the implementation of ARATs compared to non-adaptive AR training systems? (**Level 4**)

Observation protocol	Characterization (Sekaran and Bougie, 2016)	• Uncontrolled • nonparticipant observation (passive participation) • unstructured research • unconcealed (open) research
	Training theory	• Which training theory/theories is followed? • Where does the practice differ from the theory/theories?
	Participant interaction	• How much verbal interaction happens between the trainee and the trainer? What is the content? How detailed is the interaction? • How much non-verbal interaction (e.g. showing) occurs between the participants? What is the content? How detailed is it?
	Training adaption	• Do the trainers act differently with different trainees? • Which aspects of their behaviour are adapted? How?

	Assembly process	• Categorization of assembly process and product • Usage of additional tools • Costs of training products? Usable after training?
Interview protocol	Characteri-zation	• Semi-structured interview (Myers, 2009) • Face-to-face after training observation (Sekaran and Bougie, 2016)
	Participant criteria	• Are participating in the company's assembly training or did so during the last two years, either as trainer or as trainee, preferably the people observed before • Trainer: preferably at least two years experience, knows a variety of (un)finished assembly products of the company • Trainee: Has spent at least one year inside the assembly training

	Trainer interview guide	• Introduction: introduce yourself, the company you work at and your role at the company
		• Training methods:
		  ○ describe the assembly training process in detail
		  ○ what (if any) rules and guidelines are set for training
		  ○ is a specific training theory followed? If yes, which one?
		  ○ Do the guidelines deviate from the followed theory? If yes, where and why?
		  ○ Do you deviate from your company guidelines during training? If yes, where and why?
		  ○ Do you adapt your training approach to the trainee's capabilities and learning progress?
		  ○ How thoroughly do you monitor your trainees during their trial of the assembly?
		  ○ How many trainees are trained by one trainer at the same time?
		• Assembly products
		  ○ On how many different products do you train your staff to assemble in general and per person?
		  ○ How would you categorize your company's assembly/assemblies regarding complexity, economics, product traits and process traits?
		  ○ Is/Are the assembly product(s) used for training salable afterwards? Is rework needed?
		  ○ Do your training approaches differ for different products?
		• AR possibilities (after short introduction into AR systems)
		  ○ Which parts of the assembly training are you planning to assist with AR systems?
		  ○ What were your reasons to choose AR?
		  ○ Which phases could be effectively assisted through which AR system?
		  ○ Do you think the trainees would follow an AR system in the same manner as they follow you?
		  ○ Which benefits and pitfalls would you see in the implementation of AR in your assembly training?
		• Closing: Any questions or anything to add?

	Trainee interview guide	• Introduction: introduce yourself, the company you work at and your role at the company
		• Training methods
		  ○ Describe the training process you went through in detail (theoretical introduction, first guided assembly, supervised assembly, feedback on progress)
		  ○ Degree of interaction experienced and degree to which it is desired
		  ○ With how many other trainees were you sharing one trainer during the training sessions?
		• Assembly products:
		  ○ How many different products have you been trained to assemble?
		• AR possibilities (after short introduction into AR systems)
		  ○ Would you accept an AR system's authority as much as a trainer's authority?
		  ○ How would you rate AR instructions compared to trainer explanations regarding their helpfulness?
		• Closing: Any questions or anything to add?

Data analysis	Preparation (Karlsson, 2016)	• All recordings will be checked for their quality • All data will be transcribed shortly after their collection • The transcriptions will be sent for review to the participants
	Analysis	• Data analysis is of inductive nature • Results of the studies will be first summarized individually before they will be combined to achieve overall results • The data will be coded following the coding scheme of Corbin and Strauss (1990) (open, axial and selective coding) to reduce and structure the data and to display patterns resulting in a coding tree which will be compared to the literature-based expectations (Karlsson, 2016) • Inter-reliability checks will be performed on the questions answered by several people in the same case company
	Within-case synthesis: expectations	• Although generally specific methods are followed, practice deviates from those in some parts • Interaction and feedback play an important role in training and are appreciated by the trainees • In order to provide feedback and guidance, the trainee's are monitored closely during their assembly trials • Trainers change their approach for every trainee individually to fit best his capabilities and personality traits
	Cross-case synthesis	• The more valuable the product the closer the monitoring of the trainees • With increasing complexity of the product the trainees per trainer decrease • The more valuable the product is the more AR training is seen as beneficial
Guide for the case study report	The audience of the report is the research personnel of the university and the employees of the case companies. Therefore, the report has to ensure meeting academic standards regarding documentation, reliability and validity of the research and its results, but offer short and concise recommendations for practitioners.	

TABLE A.2: CASE STUDY PROTOCOL

Appendix C: Delphi Method protocol

- Desired participants: at least 2 representatives of the following groups (*categorization* according to Linstone and Turoff, 1975)
    - Augmented Reality system *experts*
    - Artificial Intelligence *experts*
    - Assembly training *stakeholders*
    - Training methods *facilitators/experts*
    - Augmented reality assembly training *experts & facilitators*
- Study questions:
    - Work out which AR system could assist which assembly phases effectively (RQ)
    - Work out the benefits and pitfalls of non-adaptive AR training systems compared to ordinary assembly training (SQ4)
    - Work out the benefits and pitfalls of ARAT systems compared to ordinary AR training systems and the conditions for those to become relevant (SQ5)
    - Work out extra requirements of ARATs compared to ordinary AR training systems (SQ6)
- Course of the study (following Schmidt, 1997):
    1. Participants will get informed about the purpose, the course and time planning of the study and about their personal rights to disagree at any point of time to any part of the study without specific reasoning
    2. First round will review the cases and conclusions from the case study, let the participants brainstorm about advantages, disadvantages and requirements and assess two exemplary constructed cases → experts shall assess the applicability of AR systems there (following Linstone and Turoff, 1975, p. 65)
    3. Similar or identical answers get grouped together by the researcher
    4. Second round is sent out where participants shall mark the most important ones out of the grouped answers from the first round; groupings are, furthermore, validated through the option to clarify his own opinion if not well-represented in given answers
    5. Grouped answers get shortlisted for the most important ones by disregarding all answers that were not selected by the majority of participants
    6. Third round gives back summaries of second round answers and opens the established shortlists for full ranking regarding their importance. Results will be compared with implications from case study.
    7. Participants will get written summary feedback of the results of the overall rankings. Furthermore they will be informed about when they will be granted access to the final report.
    8. Final report will be sent to the participants. Furthermore, personal presentations of the findings will be offered

Appendix D: Example interview transcriptions

To ensure a better reliability of the research, all interviews conducted in the case study research have been afterwards fully transcribed and verified by the participants (see section 4.1.1). The following section of an interview transcription should exemplify how the transcripts look like. **NH** as the initials of the researcher implicate the speech of the interviewer, **I** marks the words of the interviewee. In case of a specific interest, the author can also provide the full transcription of all interviews upon request.

[…]

NH: How do you control during and after the training that the people are doing their job correctly?
I: We check our employees during the training and the production occasionally. Because we are getting things to be assembled from our customers it is basically a normal production, we purchase the raw material and assemble it as the customers wants to have it. So during the production we monitor occasionally whether the products are correctly assembled or whether the people might still need some support from us. We are doing this in the production flow to control the work and give extra orders to the employees.
NH: Okay, do you train your people from new for every single product you are producing?
I: Yes, exactly. We have people that shall be appointed for a product, new people, and we have three products and for each of these products we make demonstrations with the real products. For mounting the products, you can have a look at the checklist if you like.

Going to the place where the list of steps for the assembly of a product is located.

I: This is a product that we are producing quite often and outside you can see how the product looks like in total. Here you can see the basic parts of the product which we can later teach the people when we have created the instructions. And out of this, we create the assembly instructions with a photo and a bit of text. It is not possible to have only text. Many people here are analphabetic and can only read badly or not at all, so we have to give them the information in a spoken, clear form. So we have a simple product where we want to train when new people arrive. Then we can have a look at how good the people are and then we can see if we want to put those people on different steps for a more complex lamp.
NH: Okay. Are the products you are using for the training of the people salable afterwards?
I: No, those (points at a box with disassembled desktop lamps) are a dummy order, and we assemble and disassemble those constantly for training purposes. We train basic skills with those. If someone stays longer, like three months, then we can see where we can make the best use of those people. For that, we need the buddy system where the people stay two months in one job so that one person knows the lamp well and can then explain it to other people.

[…]

Appendix E: Delphi method questionnaires

Appendix E1: First Delphi round

Introduction

Dear Participant,

Thank you for your interest and openness to participate in my research in the course of my Master thesis.

You will be participating in a Delphi Method. The purpose of this is to combine the expertise of experts in several relevant fields to new, combined knowledge. Therefore, you will be receiving in the following weeks three questionnaires, out of which this one is the first. With every new questionnaire from now, you will also see the main results of the preceding round, so the answers of you and the other experts. You are invited to take those possibly new perspectives into account while answering the next questionnaire. In this way, we are hoping that not only we as researchers profit, but also you as experts.

This test consists of overall 12 questions separated in 5 parts. Answering it will take ca. 30-45 minutes. You can at all times go back and forward in the questionnaire and also edit your answers after submission in case you forgot to mention something. Please note that the answers will only be visible to us when you submitted the questionnaire and that you cannot save your answers to continue later.

In case of any questions, you are always free to contact me as main researcher (n.heidler@student.rug.nl) or my supervisors Dr. Jos Bokhorst (j.a.c.bokhorst@rug.nl) and Dr. Adrian Small (adrian.small@ncl.ac.uk).

We are looking forward to seeing your answers.

Assessment Question

This first question has the only intention to decide whether you will be getting shorter or longer explanations with examples on the capabilities of the different systems.

How much experience do you have on Computer systems, especially Augmented Reality (AR) or artificial intelligence (AI)?

a) not that much → Leading to extensive explanations in Part 1
b) quite some → Leading to shortened explanations in Part 1

Part 1: Basic Questions – extensive explanations

Introduction to the systems: Augmented Reality (AR) systems are systems that superimpose virtual objects into the real vision of people. In most cases modern systems are spatial-aware, so they always know their position relative to the virtual object and can adapt it like in real vision. However, currently the majority of the systems are not situation-aware yet, so in the case of assembly training the system needs to be guided manually through the saved steps and it cannot provide feedback for the operator's performance.

This is where Augmented Reality Adaptive Tutors (ARATs) come into play. Based on a combination of AR technologies with Artificial Intelligence (AI), their aim is to provide an AR experience which is adapting to the environment, so i.e. the actions of its operator. This would mean for example that such a system would be able to provide feedback to the the action of the operator. It is expected that such systems might be practically usable in three to five years.

For more information on ARATs, please see https://link.springer.com/article/10.1007/s40593-014-0032-x and https://www.sciencedirect.com/science/article/pii/S0097849318301523.

The purpose of this study is to learn about the possibilities of those systems in the setting of assembly training.

In the first questions you shall compare ordinary (non-adaptive) AR training systems with a trainer educating the employees.

1. What are from your perspective the advantages of (non-adaptive) AR training systems compared to a system where a trainer educates the employees? List as many as possible.

2. What are from your perspective the disadvantages of (non-adaptive) AR training systems compared to a system where a trainer educates the employees? List as many as possible.

In the following two questions, we would like you to compare the ordinary AR training systems with the ARATs. Please consider here the system phases of development, implementation, usage and maintenance.

3. Please list from your perspective advantages that ARATs will have over ordinary (non-adaptive) AR training systems.

4. Please list from your perspective disadvantages that ARATs will have over ordinary (non-adaptive) AR training systems.

Part 1: Basic Questions – shortened explanations

Introduction to the systems: Augmented Reality (AR) systems are systems that superimpose virtual objects into the real vision of people. In most cases modern systems are spatial-aware, so they always know their position relative to the virtual object and can adapt it like in real vision. However, currently the majority of the systems are not situation-aware yet.

This is where Augmented Reality Adaptive Tutors (ARATs) come into play. Based on a combination of AR technologies with Artificial Intelligence (AI), their aim is to provide an AR experience which is adapting to the environment, so i.e. the performance of its operator.

For more information please see https://link.springer.com/article/10.1007/s40593-014-0032-x and https://www.sciencedirect.com/science/article/pii/S0097849318301523.

The purpose of this study is to learn about the possibilities of those systems in the setting of assembly training.

In the first questions you shall compare ordinary (non-adaptive) AR training systems with a trainer educating the employees.

1. What are from your perspective the advantages of (non-adaptive) AR training systems compared to a system where a trainer educates the employees? List as many as possible.

2. What are from your perspective the disadvantages of (non-adaptive) AR training systems compared to a system where a trainer educates the employees? List as many as possible.

In the following two questions, we would like you to compare the ordinary AR training systems with the ARATs. Please consider here the system phases of development, implementation, usage and maintenance.

3. Please list from your perspective advantages that ARATs will have over ordinary (non-adaptive) AR training systems.

4. Please list from your perspective disadvantages that ARATs will have over ordinary (non-adaptive) AR training systems.

Part 2: Detailed Questions

In the following two questions, we are getting into more detail in the evaluation of ARAT systems for assembly training.

5. Please briefly describe characteristics of an assembly training setting (e.g. product value, training methodology followed, ...) which you believe could be improved through the use of ARATs. Please refer to the advantages you listed in Question 3

6. What additional requirements do you think would be needed for the implementation of an ARAT system compared to a (non-adaptive) AR training system? Please think of technical, systemic, organizational and practical requirements.

Part 3: Applicability Questions

The following two questions ask about the applicability of the AR systems in different steps of the training. Please consider there the following parts from the TWI Job Instruction methodology (adapted from Dinero, 2005):

a. Prepare the worker (state job and figure out what he knows about it, get him interested and place him in correct position)

b. Present the operation (tell, show and illustrate each step clearly and completely)

c. Try-out-performance (have the worker doing the job, & correct errors, perform knowledge tests)

d. Follow-up (frequent checks of independently working worker)

7. Please shortly describe which steps of the Job instruction methodology could be improved through the use of a non-adaptive AR training system. Which performance measures (e.g. no. of mistakes) would be impacted? Why?

8. Please shortly describe which steps of the Job instruction methodology could be improved through the use of an ARAT. Which performance measures (e.g. trainer-trainee-ratio) would be impacted? Why?

Part 4: Case Questions

For the following questions, please consider the two freely constructed case companies:

> *Company 1* wants to train its assembly workers for working at the assembly line of a high volume low value product. It is looking for a quick standardized and scalable training procedure. Small product defects during training are tolerable, but after the training the employee needs to work perfectly. They are looking for a system that delivers the exact same training to all its employees and makes them ready to go to the assembly line.
>
> *Company 2* wants to train its employees in a personalized manner on the manual assembly of low volume high value products. It is important that also during the training no mistakes occur

as this would make the product unsalable. They are looking for a system to train and monitor for a long period on the actual job.

9. Both companies see the potential in implementing AR in training. Assume both non-adaptive AR systems and ARATs would be fully functional in the given setting. Which system would you advise them to choose? Why?

10. Company 1 and 2 are both currently using a trainer to teach employees but feel as though the workload for these trainers is too high as they also have production responsibilities. Aside from AR systems, can you think of any other (technological) solutions to this problem? Please briefly provide the advantages and disadvantages of your solutions compared to AR systems

11. Do you feel that you had adequate information about the example companies to provide good advice? If not, what information would you have liked and why would it have been important?

Part 5: Additional comments

12. Please kindly provide any additional comments you might have regarding this first Delphi round

Appendix E2: Second Delphi round

Introduction

Dear Participant,

We would like to thank all participants who submitted the questionnaire in the first round of the Delphi Method. We do apologize to the participants who needed longer than anticipated to complete the questionnaire, we hope and think this one should be faster to complete.

In the analysis of your responses, we summarized and grouped your responses and will now provide them back to you as list of arguments. If you do not see your answer properly represented in the descriptions or have a new idea, please use the 'Other' option.

In this questionnaire, we are aiming to narrow down the responses and determine a consensus on the most important aspects. Therefore, we ask you in the questions to give your opinion on which aspects you see as most important on that issue. You do not have to rank between those important aspects yet. The order in which all options are displayed is selected randomly. Apart from two questions, all questions are designed like that.

Overall, the questionnaire consists of ten questions in five sections.

You can at all times go back and forward in the questionnaire and also edit your answers after submission in case you want to add something. Please note that the answers will only be visible to us when you submitted the questionnaire and that you cannot save your answers to continue later.

In case of any questions, you are always free to contact me as main researcher (n.heidler@student.rug.nl) or my supervisors Dr. Jos Bokhorst (j.a.c.bokhorst@rug.nl) and Dr. Adrian Small (adrian.small@ncl.ac.uk).

We are looking forward to seeing your answers.

Part 1: Comparison of ordinary training to non-adaptive AR training

Just as a reminder, please find here again the definition of non-adaptive AR Training and ARATs:

Augmented Reality (AR) systems are systems that superimpose virtual objects into the real vision of people. In most cases modern systems are spatial-aware, so they always know their position relative to the virtual object and can adapt it like in real vision. However, currently the majority of the systems are not situation-aware yet, so in the case of assembly training the system needs to be guided manually through the saved steps and it cannot provide feedback for the operator's performance.

This is where Augmented Reality Adaptive Tutors (ARATs) come into play. Based on a combination of AR technologies with Artificial Intelligence (AI), their aim is to provide an AR experience which is adapting to the environment, so i.e. the actions of its operator. This would mean for example that such a system would be able to provide feedback to the the action of the operator. It is expected that such systems might be practically usable in three to five years

1. Below you find the identified advantages of non-adaptive AR training compared to ordinary training. Which of them would you regard as the most important ones? Please select up to nine.

a) Scalability of the training
b) Bigger flexibility regarding time, place and speed
c) Standardization of the training procedures
d) Better monitoring possibilities as instruction becomes less time consuming for the trainers
e) Better possibilities to spot flaws in educational material
f) New technologies better attract young people as trainees
g) Training is taken to a more visual, interactive level
h) Instruction possibility for many different assembly combinations without the need for extra explanation by the trainer
i) Trainers have more available resources to allow time for other responsibilities
j) Better cost-efficiency in the operational stage
k) If a system runs on common devices as smartphones, the usage of the software would be possible nearly unlimited
l) High efficiency regarding human resources
m) Other…

2. Below you find the identified disadvantages of non-adaptive AR training compared to ordinary training. Which of them would you regard as the most important ones? Please select up to seven.

a) Trainees have to motivate themselves more
b) Trainees do not get directly monitored
c) Training not personalized
d) Difficult, time-consuming and expensive set-up and development for the training
e) Acceptance of instructions by a system can be problematic
f) Knowledge of technology needed
g) Trainees do not receive personal feedback
h) Missing flexibility for small situation changes
i) Good structuring and rationalization of the training process needed
j) Reduction of human estimation, judgement and decision-power
k) Other: __________

Part 2: Comparison of ARATs to non-adaptive AR training

3. Below you find a list of the advantages of ARATs compared to non-adaptive AR training systems you identified. Which ones of them would you regard as the most important ones? Please select up to six.

a) Learning could be done "Just-in-Time", so you receive information right when you need it
b) Full monitoring through system possible
c) Personalized training
d) Improved training outcome through higher efficiency
e) More motivational for trainees
f) Broader usage possibilities also outside a training environment if needed
g) Adaptability of instructions
h) More detailed instructions
i) Systems could gather experiences to improve the teaching
j) Other: _________

4. Below you find a list of the disadvantages of ARATs compared to non-adaptive AR training systems you identified. Which ones of them would you regard as the most important ones? Please select up to four.

a) More complex in development and maintenance
b) Big datasets required
c) High initial costs
d) Acceptance of the trainees unclear
e) Lower flexibility to implement new products into the system
f) High reliability in spatial-awareness required for a stable system
g) Other: _________

Part 3:ARAT characteristics

5. Below you find the characteristics of an assembly process that are necessary for the successful implementation of ARAT, as identified in the previous questionnaire. Which of them do you regard as the most important ones? Please select up to six.

a) A detailed process of assembly steps has to be followed very precisely
b) The trainees are lower-skilled and need more guidance and monitoring
c) Fast reactions to errors are crucial
d) Possibility to clearly define which information of training shall be analyzed at what moment
e) High product complexity
f) High product value
g) The order of the assembly is crucial
h) The trainees not only need to learn the operation itself, but also how to be fast in it during their training
i) Other: ___________

6. Both, non-adaptive AR training systems and ARATs have certain requirements that need to be fulfilled so that the system works properly. Below, you find the list of identified requirements that ARATs have in addition to those non-adaptive AR training systems already have. Which of those do you see as the biggest challenges? Please select up to six.

a) Very big data sets regarding the operator's past and present performance and knowledge
b) More computing power (which by today is not possible in a portable format)
c) More sensors to create situation-awareness
d) Evaluation algorithms for all steps
e) Definitions of when, where and how feedback shall be given
f) Create awareness and trust of the trainees in the system and its functioning

g) Safety and liability responsibilities have to be settled
h) Clear definition of necessary context is necessary to not overload the system, but get the situation-awareness needed
i) Other: __________

Part 4: AR systems in TWI Job Instruction

In the following two questions, we would like to ask you for a more detailed opinion. Following on the questions of WHERE you think in the TWI Job Instruction methodology both kinds of AR can be advantageous, we are now asking you HOW it could be done.

Therefore, you will first get a briefing on the opinions and arguments mentioned in the last round. Based on that and your own opinion, please give us a broad brainstorming of how you would design the training process. This means that we are interested to hear how non-adaptive AR/ARATs and ordinary training methods can work together in the process of TWI Job Instruction.

As a short reminder, please see again the discussed methodology:

A. Prepare the worker (state job and figure out what he knows about it, get him interested and place him in correct position)
B. Present the operation (tell, show and illustrate each step clearly and completely)
C. Try-out-performance (have the worker doing the job, & correct errors, perform knowledge tests)
D. Follow-up (frequent checks of independently working worker)

Non-adaptive AR systems

In the preceding questionnaire, a consensus was reached that the first two phases (a & b) of the TWI Job Instruction methodology could be improved by non-adaptive AR while the opinions on the third phase (c) were separated. For the last phase you all agreed that non-adaptive AR would not help.

While some of you argued that the missing adaption possibilities and, therefore, the missing feedback on the try-out performances destroys the systems' chances to improve step c, others still saw possibilities there through the trainees or trainers manually clicking through the steps in order for the worker to memorize it while doing the task.

7. Based on your own opinion and the further arguments you just read, how would you design the TWI Job Instruction process with using non-adaptive AR systems? Which phase would be carried out by the system, which by the trainer? Which input would the trainer give to the system? And which data could the trainer get from the system? Where could the trainer and the system collaborate to deliver an optimal training? Please try to design a broad draft.

Augmented Reality Adaptive Tutors (ARATs)

Regarding the applicability of ARAT systems, the picture from the last questionnaire was similarly diverse. Now all of you agreed on the applicability in phases a, b and c. However, ca. 60% saw the possibility to use ARATs also in part d, so in the entire process.

For step d some of you saw good possibilities here as through the bigger amount of data on the assembly the system can improve its feedback and will, therefore, also get more valuable in following up the trainee's performance or to retrain him when he has not performed a task for a longer period. However, the usage of ARATs for the follow-up would also mean that more hardware will be needed for the company to follow-up its trainees as this happens in most occasions 'on the job' and over a long period of time.

8. Based on your own opinion and the further arguments you just read, how would you design the TWI Job Instruction process with using ARATs? Which phase would be carried out by the system, which by the trainer? Which input would the trainer give to the system? And which data could the trainer get from the system? Where could the trainer and the system collaborate to deliver an optimal training? Please try to design a broad draft.

Part 5: Cases

Finally, it is time to consider again the case companies from the preceding questionnaire. Just as a reminder, again the descriptions:

Company 1 wants to train its assembly workers for working at the assembly line of a high volume low value product. It is looking for a quick standardized and scalable training procedure. Small product defects during training are tolerable, but after the training the employee needs to work perfectly. They are looking for a system that delivers the exact same training to all its employees and makes them ready to go to the assembly line.

Company 2 wants to train its employees in a personalized manner on the manual assembly of low volume high value products. It is important that also during the training no mistakes occur as this would make the product unsalable. They are looking for a system to train and monitor for a long period on the actual job.

Furthermore, it should be stated that company 1 assembles a product of low complexity, while company 2's product has a high complexity.

The analysis of your previous answers showed that all of you had the same opinion regarding case company 2, while for company 1 the replies were mixed.

In the following two questions, you should imagine you were working for the case company and think about whether to implement non-adaptive AR or ARATs in the assembly training.

9. Below, you can find a list of arguments provided for case company 1. Some of them support the usage of non-adaptive AR for case company 1 (marked with "pro AR"), some support the usage of ARATs (marked with "pro ARAT"). Please indicate which ones of them you find the most convincing. Please select up to four.

a) Pro AR: Standardized training can be performed by non-adaptive AR
b) Pro AR: Personalization of training not needed and, therefore, too expensive
c) Pro ARAT: Companies should always strive for perfection and never allow errors, so ARATs offer better possibilities to do so
d) Pro ARAT: Adaptivity and spatial-awareness are always helpful for systems like AR
e) Pro ARAT: Price difference will be rather small when both systems are fully developed (as assumed here)
f) Other: ___________

10. Below, you find the list of different arguments you gave for case company 2. All of them are supporting the usage of ARATs in this company. Please indicate which of them you find the most convincing. Please select up to four.

a) ARATs offer the personalization needed

b) As quality is important, it should be ensured as thoroughly as possible
c) High value enhances need for perfection
d) ARATs offer possibility to train and monitor, so are useful for multiple purposes
e) ARATs give feedback to mistakes made, which helps the company to avoid them to be undetected
f) Other: __________

Appendix E3: Third Delphi round

Introduction

Dear Participant,

We would like to thank all participants who submitted the questionnaire in the second round of the Delphi Method. You are now about to start the third and final round. This round should also be the quickest for you to answer.

In the analysis of the last round, we looked at your answers and combined them into a consolidated list of aspects that the majority of you considered as important. We now want you to rank those aspects regarding their importance. This will result in a ranked list of aspects that companies will be able to reference when considering the implementation of AR into their assembly training.

The questionnaire consists of eleven questions separated in five sections.

You can at all times go back and forward in the questionnaire and also edit your answers after submission in case you want to add something. Please note that the answers will only be visible to us when you submitted the questionnaire and that you cannot save your answers to continue later.

In case of any questions, you are always free to contact me as main researcher (n.heidler@student.rug.nl) or my supervisors Dr. Jos Bokhorst (j.a.c.bokhorst@rug.nl) and Dr. Adrian Small (adrian.small@ncl.ac.uk).

We are looking forward to seeing your answers.

Part 1: Comparison of ordinary training to non-adaptive AR training

1. Below are the six advantages of non-adaptive AR training compared to ordinary training methods that were identified as most important in the previous round. Please rank these advantages according to their importance on an ordinal scale with 1 being the most important and 6 the least important.

a) Scalability of the training
b) Standardization of the training procedures
c) Instruction possibility for many different assembly combinations without the need for extra explanation by the trainer
d) Trainers have more available resources to allow time for other responsibilities
e) Bigger flexibility regarding time, place and speed
f) High efficiency regarding human resources

2. Below are the four disadvantages of non-adaptive AR training compared to ordinary training methods that were identified as most important in the previous round. Please rank these advantages according to their importance on an ordinal scale with 1 being the most important and 4 the least important.

a) Training is not personalized
b) Trainees do not receive personal feedback

c) Missing flexibility for small situation changes
d) Difficult, time-consuming and expensive set-up and development for the training

Part 2: Comparison of ARATs to non-adaptive AR training

Just as a reminder, please find here again the definitions of non-adaptive AR Training and ARATs:

Augmented Reality (AR) systems are systems that superimpose virtual objects into the real vision of people. In most cases modern systems are spatial-aware, so they always know their position relative to the virtual object and can adapt it like in real vision. However, currently the majority of the systems are not situation-aware yet, so in the case of assembly training the system needs to be guided manually through the saved steps and it cannot provide feedback for the operator's performance.

This is where Augmented Reality Adaptive Tutors (ARATs) come into play. Based on a combination of AR technologies with Artificial Intelligence (AI), their aim is to provide an AR experience which is adapting to the environment, so i.e. the actions of its operator. This would mean for example that such a system would be able to provide feedback to the the action of the operator. It is expected that such systems might be practically usable in three to five years

3. Below are the five advantages of ARATs compared to non-adaptive AR training systems that were identified as most important in the previous round. Please rank these advantages according to their importance on an ordinal scale with 1 being the most important and 5 the least important.

a) Learning could be done "Just-in-Time", so you receive information right when you need it
b) Personalized training
c) Adaptability of instructions
d) Improved training outcome through higher efficiency
e) Systems could gather experiences to improve the teaching

4. Below are the three disadvantages of ARATs compared to non-adaptive AR training systems that were identified as most important in the previous round. Please rank these advantages according to their importance on an ordinal scale with 1 being the most important and 3 the least important.

a) More complex in development and maintenance
b) Big datasets required
c) High initial costs

Part 3: ARAT characteristics

Next, we are looking at the characteristics an assembly process should have in order to be a good environment for an ARAT to be implemented. Furthermore, we are interested in the perceived biggest challenges of implementing ARAT in addition to those of a non-adaptive AR system.

5. Below are the most important characteristics of an assembly process necessary for the successful implementation of an ARAT training system, as identified in the previous round. Please rank them according to their importance on an ordinal scale with 1 being the most important and 2 the least important.

a) Fast reactions to errors are crucial
b) High product complexity

6. Both, non-adaptive AR training systems and ARATs have certain requirements that need to be fulfilled so that the system works properly. Below are the four most challenging requirements that ARATs have in addition to those non-adaptive AR training systems already have as identified in the

previous round. Please rank them on an ordinal scale with 1 being the biggest challenge and 4 the smallest of those options.

a) Very big data sets regarding the operator's past and present performance and knowledge
b) Evaluation algorithms for all steps
c) Definitions of when, where and how feedback shall be given
d) More sensors to create situation-awareness

Part 4: AR systems in TWI Job Instruction

In the following questions, we are interested in how your opinion has developed after learning about each other's arguments on how to represent the steps of TWI Job Instruction with AR systems. Furthermore we are looking at to which extent the trainers will still have to be included in those. Therefore, after a short wrap-up of the solutions proposed in the last round, we are interested in what you would choose for each step if you had to take the decision for only this single step and if you had to take it for all steps together.

As a short reminder, please see again the discussed methodology:

A. Prepare the worker (state job and figure out what he knows about it, get him interested and place him in correct position)
B. Present the operation (tell, show and illustrate each step clearly and completely)
C. Try-out-performance (have the worker doing the job, & correct errors, perform knowledge tests)
D. Follow-up (frequent checks of independently working worker)

Through your responses from the last round, we were able to get a better image of how the different training methods could be implemented in the different steps of the TWI process. Below is a summary of your design ideas for each step in the process. Please consider this description when answering the following questions.

Step A could be designed in non-adaptive AR in such a way that it gives a design preview of the assembly steps and some quiz questions. The input for this would have to come from a trainer. With an ARAT, the system would still need input from about the trainee, but could already during this phase start gathering data for the later phases. However, also an ordinary trainer was considered as best option to introduce the trainee into his task and how to handle the AR system.

For *Step B*, non-adaptive AR could present the instructions and further explanations. An ARAT could also record the necessary steps in advance and replay them afterwards to the trainee after a check for correctness by the trainer.

In *Step C*, a non-adaptive AR system could guide the trainees through the assembly, but the trainers would still have to watch out for mistakes to correct and the guidance would be standardized. An ARAT could however give the guidance to the trainees "Just-in-Time", so only and right at the moment they need it. Furthermore, the system would learn through every operation it is included and could give information about the trainee's performance to the trainers.

In *Step D*, non-adaptive AR systems could be used as solution to remind the trainees of the correct assembly, so some kind of target-performance comparison by themselves. Objective revisions or checks would have to be performed by the trainers. An ARAT however could be also used as objective monitoring and Just-in-Time mistake correction measure and could also collect much data on the operations during this usage to further improve its database.

We also received the opinion that for complex tasks training necessarily needs to be done by trainers and AR technologies might be then helpful in the follow-up.

7.1 Please consider the situation a company would be interested in taking a decision for ONLY ONE part of the TWI Job Instruction methodology. Which solution would you recommend them to choose for each step without looking at the other steps?

	Ordinary training methods	non-adaptive AR system	ARAT
Step A: Prepare the worker	◻	◻	◻
Step B: Present the operation	◻	◻	◻
Step C: Try-out performance	◻	◻	◻
Step D: Follow-up	◻	◻	◻

7.2 In your previously selected choices, how much trainer inclusion would be needed in the steps?

	High trainer inclusion	Low trainer inclusion
Step A: Prepare the worker	◻	◻
Step B: Present the operation	◻	◻
Step C: Try-out performance	◻	◻
Step D: Follow-up	◻	◻

8.1 Please consider now that a company is building up an entire system following the TWI Job Instruction methodology. Which solution would you choose now for each step taking into account the connectivity between the steps?

	Ordinary training methods	non-adaptive AR system	ARAT
Step A: Prepare the worker	◻	◻	◻
Step B: Present the operation	◻	◻	◻
Step C: Try-out performance	◻	◻	◻
Step D: Follow-up	◻	◻	◻

8.2 In your previously selected choices, how much trainer inclusion would be needed in the steps?

	High trainer inclusion	Low trainer inclusion
Step A: Prepare the worker	◻	◻
Step B: Present the operation	◻	◻
Step C: Try-out performance	◻	◻
Step D: Follow-up	◻	◻

Part 5: Cases

Finally, it is time to consider again the case companies from the preceding questionnaire. Just as a reminder, again the descriptions:

Company 1 wants to train its assembly workers for working at the assembly line of a high volume low value product of low complexity. It is looking for a quick standardized and scalable training procedure. Small product defects during training are tolerable, but after the training the employee needs to work perfectly. They are looking for a system that delivers the exact same training to all its employees and makes them ready to go to the assembly line.

Company 2 wants to train its employees in a personalized manner on the manual assembly of low volume high value products of high complexity. It is important that also during the training no mistakes occur as this would make the product unsalable. They are looking for a system to train and monitor for a long period on the actual job.

In the following three questions, we would like you to rank the most important arguments and reconsider the recommendation for case company 1 whether they should implement non-adaptive AR training systems or ARATs.

9. Below you find the two arguments regarding case company 1 a majority of you identified as important. As it turned out, both of them supported the use of non-adaptive AR while in the first questionnaire the majority supported the use of ARATs. Please rank them regarding their importance on an ordinal scale with 1 being the most important and 2 the least important.

a) Standardized training can be performed by non-adaptive AR
b) Personalization of training not needed and, therefore, too expensive

10. After considering the above arguments, what course of action would you recommend for Case Company 1?

- ☐ Implement a non-adaptive AR system
- ☐ Implement an ARAT system

11. Below are the two most important arguments supporting the use of ARATs in Case Company 2, as identified in the previous round. Please rank them according to their importance on an ordinal scale with 1 being the most important and 2 the least important.

a) As quality is important, it should be ensured as thoroughly as possible
b) ARATs give feedback to mistakes made, which helps the company to avoid them to be undetected

As you all agreed regarding the choice of system in the first questionnaire, it would not add any value to reconsider that question again.

Appendix F: Kendall's coefficient of concordance w

In the following section, the formula used to calculate Kendall's w as a measure of the degree of consensus of the rankings in the final Delphi round is given.

$$w = \frac{\sum_{i=1}^{N}(\bar{R_i} - \bar{R})^2}{\frac{N(N^2-1)}{12}}$$

with k = number of judges
N = number of objects being ranked
$\bar{R_i}$ = average of the ranks assigned to the ith object
$\bar{R}$ = average of the ranks assigned across all objects (Siegel, 1988)

Appendix G: Delphi groupings

In the following section, all aspects mentioned by the experts in the first round are given grouped form. Furthermore, the results on the importance vote of those aspects performed in the second round are given. Those formed the basis for the shortlists the experts ranked in the third round.

Advantage	Vote	Disadvantage	Vote
High efficiency regarding human resources	55 %	Trainees have to motivate themselves more	44 %
Bigger flexibility regarding time, place and speed	55 %	Trainees do not get directly monitored	22 %
Standardization of the training procedures	66 %	Training not personalized	66 %
Better monitoring possibilities as instruction becomes less time consuming for the trainers	33 %	Difficult, time-consuming and expensive set-up and development for the training	55 %
Better possibilities to spot flaws in educational material	44 %	Acceptance of instructions by a system can be problematic	44 %
New technologies better attract young people as trainees	33 %	Knowledge of technology needed	33 %
Training is taken to a more visual, interactive level	44 %	Trainees do not receive personal feedback	66 %
Instruction possibility for many different assembly combinations without the need for extra explanation by the trainer	77 %	Missing flexibility for small situation changes	55 %
Trainers have more available resources to allow time for other responsibilities	66 %	Good structuring and rationalization of the training process needed	33 %
Better cost-efficiency in the operational stage	33 %	Reduction of human estimation, judgement and decision-power	33 %
If a system runs on common devices as smartphones, the usage of the software would be possible nearly unlimited	22 %		
Scalability of the training	66 %		

TABLE A.3: GROUPED ADVANTAGES AND DISADVANTAGES OF NON-ADAPTIVE AR TRAINING COMPARED TO ORDINARY TRAINING METHODS

Advantage	Vote	Disadvantage	Vote
Learning could be done "Just-in-Time", so you receive information right when you need it	66 %	More complex in development and maintenance	88 %
Full monitoring through system possible	44 %	Big datasets required	66 %
Personalized training	66 %	High initial costs	55 %
Improved training outcome through higher efficiency	55 %	Acceptance of the trainees unclear	33 %
Systems could gather experiences to improve the teaching	55 %	High reliability in spatial-awareness required for a stable system	44 %
Broader usage possibilities also outside a training environment if needed	22 %	Lower flexibility to implement new products into the system	33 %
Adaptability of instructions	66 %		
More detailed instructions	33 %		
More motivational for trainees	44 %		

TABLE A.4: GROUPED ADVANTAGES AND DISADVANTAGES OF ARATS COMPARED TO NON-ADAPTIVE AR TRAINING

Characteristic	Vote
A detailed process of assembly steps has to be followed very precisely	44 %
The trainees are lower-skilled and need more guidance and monitoring	33 %
Fast reactions to errors are crucial	66 %
Possibility to clearly define which information of training shall be analysed at what moment	44 %
High product complexity	66 %
High product value	33 %
The order of the assembly is crucial	33 %
The trainees not only need to learn the operation itself, but also how to be fast in it during their training	33 %

TABLE A.5: GROUPED ASSEMBLY CHARACTERISTICS NECESSARY FOR AN ARAT IMPLEMENTATION

Requirement	Vote
Very big data sets regarding the operator's past and present performance and knowledge	66 %
More computing power (which by today is not possible in a portable format)	22 %
More sensors to create situation-awareness	55 %
Evaluation algorithms for all steps	66 %
Definitions of when, where and how feedback shall be given	66 %
Create awareness and trust of the trainees in the system and its functioning	33 %
Safety and liability responsibilities have to be settled	22 %
Clear definition of context is necessary to not overload the system, but get the situation-awareness needed	44 %

TABLE A.6: GROUPED EXTRA REQUIREMENTS OF ARATS COMPARED TO NON-ADAPTIVE AR TRAINING SYSTEMS

Pro non-adaptive AR	Vote	Pro ARAT	Vote
Standardized training can be performed by non-adaptive AR	77 %	Companies should always strive for perfection and never allow errors, so ARATs offer better possibilities to do so	22 %
Personalization of training not needed and, therefore, too expensive	66 %	Adaptability and spatial-awareness are always helpful for systems like AR	44 %
		Price difference will be rather small when both systems are fully developed (as assumed here)	44 %

TABLE A.7: GROUPED ARGUMENTS CASE COMPANY A (SEE TABLE 5.8)

Pro ARAT	Vote
ARATs offer the personalization needed	33 %
As quality is important, it should be ensured as thoroughly as possible	88 %
High value enhances need for perfection	44 %
ARATs offer possibility to train and monitor, so are useful for multiple purposes	44 %
ARATs give feedback to mistakes made, which helps the company to avoid them to be undetected	88 %

TABLE A.8: GROUPED ARGUMENTS CASE COMPANY B (SEE TABLE 5.8)

Appendix H: Degree of consensus amongst the Delphi expert groups

In the following, the values of Kendall's w for the rankings given in Table 5.3 to Table 5.5 are examined more thoroughly regarding the different groups of Delphi experts that responded the last rounds questionnaire. For the group of AR experts with a k=2, a critical value for an α=0,05 significance cannot be reached. The same holds true for k-3 with N<4 aspects ranked.

Furthermore, one training expert participated in the last round, which was left out here.

Question	N	Overall (k=9)		AI experts (k=3)		AR experts (k=2)		Assembly train. exp. (k=3)	
		w	Crit. value	w	Crit. value	w	Crit. value	w	Crit. value
Advantages non-adaptive AR (Table 5.3)	6	0,254*	0,241	0,352	0,660	0,371	/	0,340	0,660
Disadvantages non-adaptive AR (Table 5.3)	4	0,012	0,287	0,111	/	0,100	/	0,111	/
Advantages ARAT (Table 5.4)	5	0,064	0,259	0,089	0,716	0,350	/	0,511	0,716
Disadvantages ARAT Table 5.4)	3	0,259	0,333	0,333	/	0,250	/	0,333	/
Extra requirements (Table 5.5)	4	0,131	0,287	0,200	/	0,100	/	0,200	/

TABLE A.9: KENDALL'S W PER EXPERT GROUP

Appendix I: Decision Aid for companies

In the following section, the recommended procedure to make use of the results of this research for companies planning to implement AR in their assembly training will be described briefly in the following figure. For a more thorough evaluation the tables from Appendix G may be also considered.

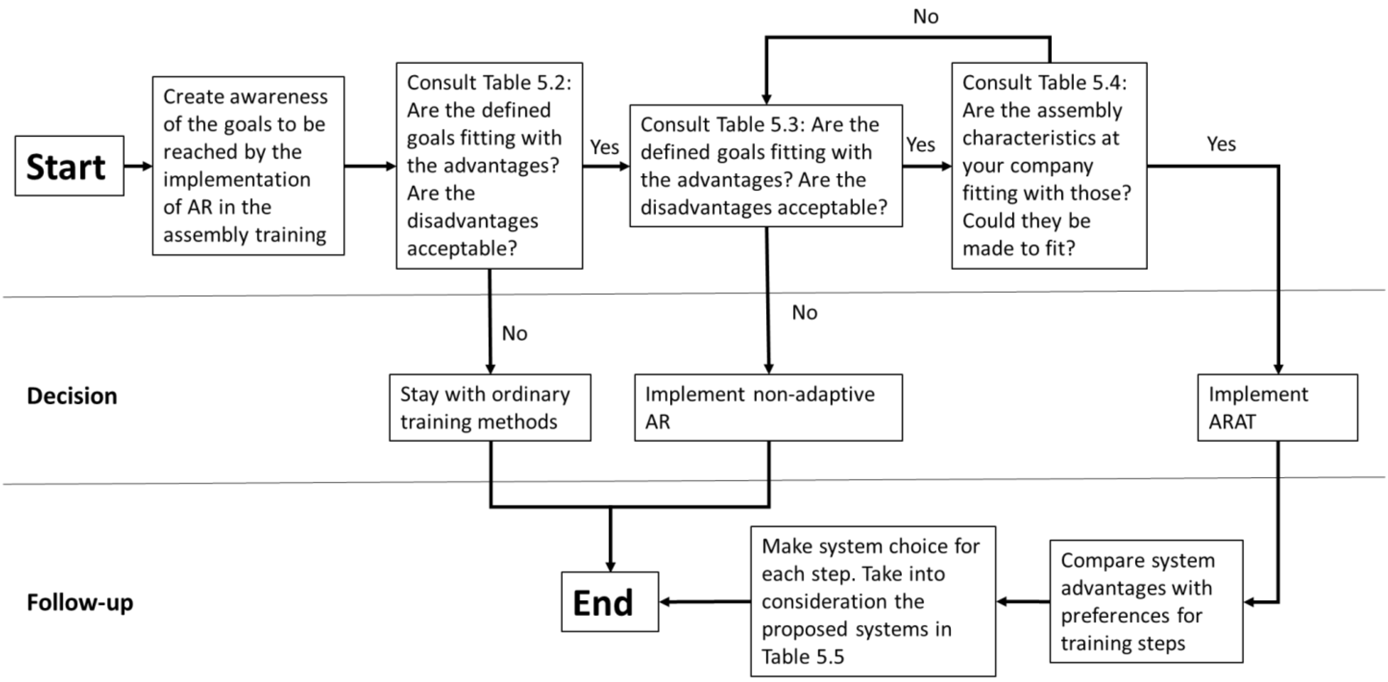

FIGURE A.1: DECISION AID FOR COMPANIES