Designing an Augmented Reality Authoring Tool to Support Complex Tasks. A Design Science Study Using Cognitive Load Theory

Kay Hönemann(✉), Björn Konopka, and Manuel Wiesche

TU Dortmund University, Dortmund, Germany
{kay.hoenemann,bjoern.konopka,manuel.wiesche}@tu-dortmund.de

Abstract. Despite the potential of augmented reality (AR) to support and guide complex industrial tasks, the technology is still not broadly applied. One possible reason for this is that the creation of AR content is highly complex and requires programming skills and deep spatial knowledge. AR authoring tools can help address this complexity by enabling non-developers to create self-sufficient AR content. Therefore, this paper proposes a theory-driven design for AR authoring tools that allows non-developers to create self-sufficient AR-based instructions to support complex tasks. Based on ten interviews with experts working with AR authoring tools and a following focus group with eight participants, we propose three design principles for future AR authoring tools in the engineering context. These design principles are instantiated in two prototypes of different richness and evaluated in an experiment with 23 students. Our study shows that the cognitive load is slightly increased when using the extensive AR authoring tool, but it also shows that significantly better results can be achieved with the extensive AR authoring tool. We contribute by providing design principles for AR authoring tools for creating AR-based instructions, which extend the existing AR authoring research in the industrial context.

Keywords: Augmented Reality · AR Authoring · Design Science Research · Cognitive Load Theory

1 Introduction

We often find ourselves in situations where we encounter problems in our daily work. When we are not able to handle the issues on our own, we either seek help from people with more experience and expertise [1], or we are supported by specific technologies that help us to handle the problem and at the same time ensure that we complete our work in a compliant manner [2]. Especially in complex work environments and in the execution of knowledge-intensive tasks, these technologies are needed. The tasks include activities that, on the one hand, have a strong contextual reference, influenced by several environmental factors. On the other hand, the task requires an enormous amount of knowledge. An example is the complexity of designing and developing new information systems [3]. Another example would be the use of AR in the safety-critical service process of soil

A. Gerber and R. Baskerville (Eds.): DESRIST 2023, LNCS 13873, pp. 87–101, 2023.
https://doi.org/10.1007/978-3-031-32808-4_6

sounding [4] or the maintenance and repair of technical assets [5]. For instance, in software engineering, users are supported by a specific technology during the development phase to improve process compliance [6]. In recent years, digitally enriched guidance has increased significantly through augmented reality (AR) technology for these tasks. Very early on, researchers and practitioners recognized the enormous potential of these digitally enriched guidance systems [7]. Yet, the technology has mainly been applied in the engineering context only in prototype form. One reason could be that creating AR content is very complex and requires strong programming skills and deep spatial knowledge [8]. Consequently, 64% of all AR applications in the engineering sector are individual developments [9]. Individual developments are often not very practical for SMEs, as even small changes, which are part of the daily work for such volatile things as instructions, mean significant development costs.

AR authoring tools can help address this challenge. Authoring tools are software programs that allow users to create and publish AR experiences on different platforms on different types of hardware [8, 10]. However, a closer look at the existing tools reveals that they are still used by experienced software developers, as some of the more complex functions require extensive programming knowledge and 3D modeling skills [8]. Therefore, the creation of AR content by non-developers remains a challenge.

Many AR studies consider how the ideal work instruction looks like in augmented reality. For example, one study examines what information is needed to depict an instruction in AR [11]. Even the effectiveness of AR guidance compared to the usual paper guidance has been researched in recent years [12]. Other studies, in turn, have explored which type of AR representations are most suitable for AR guidance [13]. In contrast, the creation of AR content, so-called AR authoring, is still an unknown area of research. These findings refer exclusively to the use of AR content and do not provide any information about which AR content has which kind of impact on users during the creation process. For example, in the literature, it was determined that complex visualization elements are better suited for using AR instructions [13]. However, it might be possible that the users find it more difficult to create these complex visualization elements, which could then influence the quality of the AR instructions. To the best of our knowledge, there is no study on how the complexity of AR visualizations affects the mental and cognitive abilities of the creator during the creation process. Therefore, we aim to answer the following research question: "*How does the richness of AR visualizations affect task performance, ease of use, and mental and cognitive load during the creation process?*".

To answer the research question, we follow a comprehensive Design Science Research (DSR) methodology [14] focusing on designing innovative artifacts. The research is structured as follows. The conceptual background and related work are considered in the next chapter. The third chapter gives an overview of the DSR approach, which was applied in the study. The fourth chapter describes the design principles, the software artifact, and the evaluation. Finally, in the last chapter, the practical and theoretical contribution paper is discussed, as well as the limitations and future research.

2 Conceptual Background and Related Work

2.1 AR Authoring

Despite the broad application of AR in recent years in many fields, there are still several challenges in creating AR content. Ashtari et al. identified eight key barriers defining the complexity of AR content creation [15]. These range from the initial difficulties in choosing suitable tools to the lack of clear design guidelines to the challenges in testing and evaluating AR applications. With AR authoring tools, these barriers should be overcome in the future. Nebeling & Speicher analyzed existing AR/VR authoring tools in their research. As a result, they could divide the existing tools into five groups [8]. In the first group are tools like InVision or Sketch, non-developers can use them, but they do not offer an AR/VR interface yet. Tools such as Unity or Unreal are assigned to the fifth group, where individual customizations with these tools require strong programming skills and deep spatial knowledge. In the second group, tools are assigned which can be used exactly like in the first group by non-developers. Tools in this class are no-code/low-code tools for developing AR content. They differ from conventional no-code/low-code tools from software development because they provide an AR interface allowing 3D content to be anchored in the real environment. This aspect makes the AR authoring tools unique.

One of the first examples is the AR authoring tool DART from 2005 [16]. This tool allows users to create complex relationships between the real and virtual worlds. This first approach led to new tools for different application areas to create AR content to support context-intensive tasks [17, 18]. Another important research area is AR authoring tools, which are applied to Head Mounted Displays (HMD). For example, non-developers can use tools such as HoloWFM [19] to create workflows that increase the efficiency of task execution.

2.2 AR Tools Supporting Complex Tasks

Due to the continuous implementation of smart connected products, more and more data is available to users. This rising flood of data leads to a fundamental disconnection from the real environment [20]. Thus, some researchers argue that while the real environment is three-dimensional, the data we use daily to make new decisions are trapped in two-dimensional spaces like displays [20]. The gap between the real and virtual environment hinders the ability to make the best possible decisions [20]. Researchers and practitioners are already well aware of this gap, which is why one of the biggest use cases for AR is to provide process guidance based on data directly in situ [21]. Kortekamp et al. identified that this use case is the most frequently applied in the manufacturing environment [22]. In this use case, information is displayed to the users in situ, which is required to work in a compliant manner. For example, service technicians need to know the sequence of maintenance steps and how to perform them.

AR is often used for complex tasks in the field of aerospace technology to complete the work in a process-compliant manner. For example, authors Chen et al. initiated and evaluated an AR application that guides service technicians through the assembly process of cables in large spacecraft components [23]. In another AR application, service technicians are guided through installing an aircraft's wiring harness. Using AR reduced the time required to identify the parts to be installed. As a result, maintenance time was reduced by 90% [7]. These are just two examples of how AR can be used to support complex tasks. Many more AR applications in this context exist in the literature [5, 12]. In addition to the AR applications, researchers have identified six information types (i.e., identity, location, way-to, notification, order, and orientation) with which a maintenance manual can be fully represented in AR [11].

2.3 Cognitive Load Theory

For this research, we draw on a kernel theory concerned with conceptualizing and representing design knowledge. This research draws on Cognitive Load Theory (CLT), which originates in educational psychology [24]. The basic assumption of CLT is that individuals are limited in their ability to absorb and process information during the learning process. Furthermore, the CLT framework provides guidance on how to structure information to achieve better learning outcomes [25]. We argue that using new technologies (i.e., AR authoring tool) to achieve a specific goal (i.e., creating an AR instruction) is a learning process. As the design principles must have a long-term learning impact on the users, we use the CLT as the kernel theory.

The cognitive impact of AR instructions during use is frequently studied in the literature. For example, early research has shown that AR assembly instructions in the user's field of view require fewer cognitive resources than traditional assembly instructions on display [12]. This is just one example. Recent studies have also highlighted the potential of AR instructions to reduce cognitive load [26]. However, creating AR instructions using AR authoring tools has not yet been considered from a cognitive load perspective.

3 Methodology

Our study is part of a large ongoing DSR project that provides an innovative solution to a real-world problem [27]. More specifically, we address the lack of design knowledge about how non-developers can create AR content. In this way, we want to improve SMEs' access to AR technology. We adapted Küchler & Vaishnavi's [14] approach. Therefore, we separated the overall DSR project into three successive design cycles. The study focuses on the second design cycle based on the first design cycle. It forms the basis for the final design cycle. The following section briefly describes the entire DSR project to provide additional information and highlight the overall research goal. We started the first design cycle by investigating the challenges related to task execution in the engineering context, for which we conducted a first literature review. The findings highlight that in the literature, AR applications have been predominantly instantiated and evaluated in various contexts. Many of these studies point to an enormous savings potential of AR [5, 12]. The literature also showed that despite the existing AR authoring tools, there are still enormous difficulties in creating AR content [8]. There is a significant need for research to understand better the needs of the different types of users trying to get into AR development [15]. Therefore, we conducted an evaluation study with an industrial plant supplier with a subsequent focus group. For instance, we found out that the use of HMDs is prototypically excellent for these types of tasks, but in practice, these devices are not yet suitable for daily operation due to various boundary conditions (i.e., the already dark operating environments become much darker with an HMD) [28]. The literature review and evaluation study findings form our initial meta-requirements. Based on these, we developed our first software prototype and evaluated it in a field study. In the second design cycle, we adapted the design principles based on the findings of the first design cycle and were thus able to instantiate a first software prototype and subsequently evaluate it in a laboratory experiment. In the third design cycle, we intend to replicate the results of the second design cycle in a real application context. In doing so, we want to respond to the request of Peffers et al. to evaluate DSR artifacts in more real environments [29].

4 Design Science Research Project

4.1 Problem Awareness

To further deepen our understanding of the problem, we conducted ten semi-structured interviews with experts who work with AR authoring tools in their daily work. We also conducted a focus group with eight participants to triangulate and validate the findings from the interviews. Three AR authoring experts and five users from the identified target group (i.e., service technicians) participated in the focus group. The focus group is used to gain a better understanding of the results and to facilitate and further enrich the results. The experience of the individuals with AR authoring tools differed, so 30% could be classified as beginners, 40% as intermediates, and 30% as specialists.

In line with the literature, existing AR authoring tools are empowering for AR authoring experts, but in practice, these are not used by non-developers [8]. Their justification for this is straightforward. The fault lies with the user. For example, one participant said: *'[I]… is some kind of painting tool that is so simple that customers can use it to create AR experiences without any additional help. [I] I find the solution very fast and find it correspondingly interesting that people still don't jump on it. [I] The fact that the authoring is not done by the customers themselves is due to the customers and their lack of capacity, not due to the tool'*. Another important aspect some participants mentioned was that the existing tools are designed for a specific application domain. In their opinion, this makes sense since the requirements for a marketing application differ from those for an engineering application. '*[IV]… is specialized in maintenance/service instructions. For marketing applications, it is not suitable because special requirements like font size, font, special image size are not feasible.*'. Some participants also mentioned requirements in AR authoring from which design suggestions emerged. For example, almost all participants agreed that the most time-consuming and biggest challenge for an AR authoring tool user is to obtain the correct data (i.e., CAD models) and then compress them [30].

Another challenging aspect of AR authoring is choosing the appropriate tracking method and its implementation. Tracking is unique for AR applications in this form so far. Thus, it can be assumed that non-developers have no experience in this area, but it forms the basis of nearly every AR application. Without tracking, virtual objects cannot be anchored in the user's real environment [28]. Making decisions about AR hardware also causes challenges for users. Although the interviews revealed several advantages of AR authoring tools, they also highlighted the complexity of the approach. Especially topics like tracking or handling 3D data are complex and require some prior experience. Therefore, future AR authoring tools must guide non-developers through these processes.

4.2 Suggestion

We obtained substantial evidence about the design of AR authoring tools from the industrial context from the interviews. The first meta-requirement (**MR1**) refers to the effective use of a system, which results from the co-dependence between the system and the task. To counteract this representative complexity, the system should be designed so that the user is supported in the execution of the task and at the same time that only a few semantic understandings are necessary for the execution of the task [31]. The authors vom Brocke et al. propose ten principles of how good process management can be implemented in practice. First, the system should be designed to fit into the organizational context on the one hand, and on the other, the system should make opportune use of AR technology [32]. The creation of AR-based instructions differs significantly from the creation of conventional instructions, as digital content is anchored in the user's real environment [33]. Users must therefore be able to create both two-dimensional and three-dimensional processes. Our first design principle (DP), which we propose, is based on this meta-requirement: **DP1**: *Provide the system with various process steps in order to clearly separate the creation of pure 2D and 3D content for non-developer.*

The second meta-requirement (**MR2**) refers to a library of 2D and 3D elements to completely represent an instruction in AR. The authors Gattullo et al. identified six types of information that can be used to represent work instructions in AR. For this reason, the users must be able to access different 2D and 3D elements [11]. Both distinguishing intrinsic and extraneous cognitive load are related to the difference between novices and experts in their learning process [25]. These effects imply that novices need more detailed information about a system to be able to use it to overcome the knowledge gap. On the other hand, experts with a greater knowledge level in using AR draw on their prior knowledge. Since users use an AR authoring tool with different levels of prior knowledge, the system must provide a set of AR elements requiring different information levels to create AR instructions of different complexities [34]. The third meta-requirement (**MR3**) relates to the ability of users to create their media. The decision between using complex and abstract AR representations is a matter of disagreement among researchers. Some studies indicate that using complex AR representations distracts users, negatively affecting the error rate [35]. On the other hand, some studies demonstrate that much better results can be achieved with complex AR representations. However, these also show that similarly good results can be achieved with abstract AR representations in combination with media like photos or videos [13]. The two meta requirements described above form the foundation for the second design principle we propose: **DP2**: *Provide the system with a library of display and information elements which requires different information levels so that non-developers can entirely visualize instruction in Augmented reality of different complexities.*

The fourth meta-requirement (**MR4**) refers to the general design approach regarding the target group, the non-developers. On the one hand, just-in-time access to information during task performance reduces cognitive load, and on the other hand, it enables learning [36]. With augmented reality instructions, only the required information is displayed in the right place in the real environment at the right time. The non-developers are empowered to create self-sufficient AR content to increase their learning outcomes sustainably [25]. In order for users to adopt new interactive technologies, they must be able to adapt the technologies to their practice [37]. This perspective leads to discourse in research about the end-user development approach. This is defined as various methods, techniques, and tools that allow non-developers to create, adapt and extend the software for themselves [38]. According to Nebeling and Speicher, these tools belong to the second group of AR authoring tools [8]. The fifth meta-requirement (**MR5**) relates to the environment of the industrial context. Asset maintenance often occurs at the customer's site in inaccessible locations. Many studies use HMDs for this application context and show the great potential associated with them [7, 39]. Nevertheless, the hardware is not yet mature enough for daily use since most HMDs are heavy and have short battery life and a limited field of view [40]. For this reason, the AR authoring tool must be used on application-grade hardware. The two meta requirements thus form our final design principle that we propose: **DP3**: *Provide the system with an end-user development and just-in-time access to information approach empowering non-developers with different levels of experience to create self-sufficient Augmented Reality content for a specific application domain.* By initiating the prototypes based on all these approaches, we intend to evaluate these approaches and then adapt the design principles accordingly.

4.3 Development

To instantiate our design principles, we developed an AR authoring tool prototype. The first design principle specifically maps onto the possibility of creating a process flow for the instructions. The instructions' structure and sequence are determined in the so-called Node-Editor. Here three different node types are available to users. The first node type is the Info-Node, where only 2D elements (i.e., text) can be added. An Instruction-Node reflects one step in an instruction manual, and an Exploration-Node, where only location-dependent content can be displayed. The second design principle specifically maps onto the display and information elements used to represent the instructions in AR. For this purpose, a library of seven annotations is available to the users, which can be anchored in the users' real environment. These range from simple arrows, which refer to a position in the real environment, to complex tethers, which act as navigation through the real environment. In addition to the 3D elements, 2D elements such as media or text can be added to an instruction step. The third design principle maps onto the foundational design of a standalone AR authoring tool that operates on a tablet without installing additional software or plugins or deploying additional hardware. The tool can be used to create AR instructions as well as to support the execution of complex tasks. Through the use of an end-user development approach, the Graphical User Interface does not allow users to implement their code. Instead, users can create instructions by drag and drop and enrich them with 2D and 3D elements. Figure 1 shows the node editor on the left and the authoring environment of the AR authoring tool on the right.

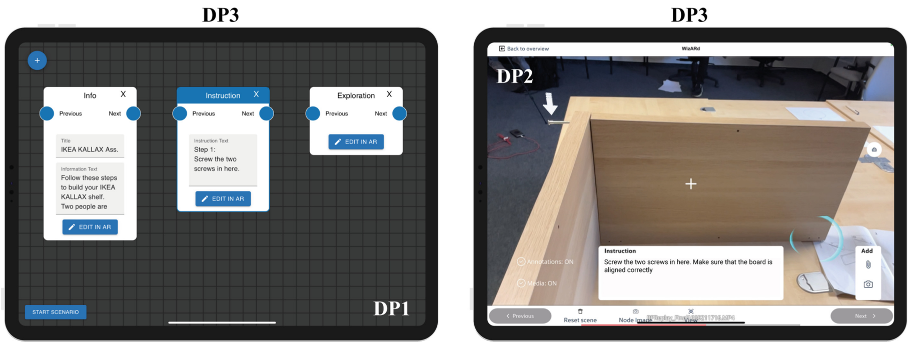

Fig. 1. Software Prototype: AR Authoring Tool

Since participants' preferences varied widely and there is disagreement in the existing literature about the complexity of AR visualizations, we developed a second prototype with different AR visualizations (DP2). Based on the evaluation results, we want to identify the most important display and information elements and adapt our design principles accordingly. As described above, the first prototype provides users with a total of seven different annotations. In the second prototype, only two annotations are available to users (i.e., arrow and point of interest). Furthermore, users do not have the option to add media to an instruction step. Figure 2 shows the AR authoring environment of both prototypes. The right side shows the prototype with extensive visualization elements (i.e., various abstract AR elements, including media), and the left side shows the prototype with reduced visualization elements (i.e., a few abstract AR elements without media).

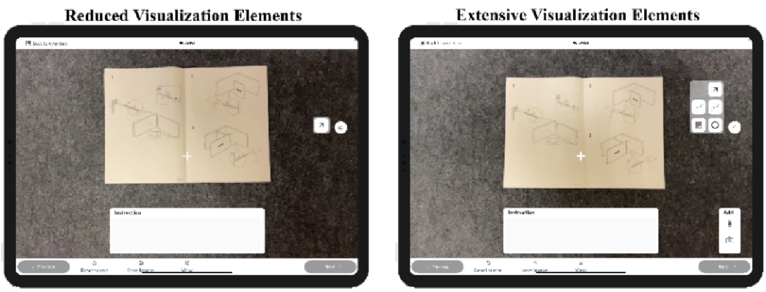

Fig. 2. AR Authoring environment of both prototypes

4.4 Evaluation

To examine task performance, we examined the effectiveness, efficiency, and quality of the created AR instruction as key performance indicators. We also analyzed the mental effort, cognitive load, and ease of use to assess important performance indicators related to the CLT. The following hypotheses regarding task performance and the CLT when using the AR authoring tool were investigated in this experiment.

- **H1a:** Increased richness of the tool positively affects effectiveness.
- **H1b:** Increased richness of the tool positively affects efficiency.
- **H1c:** Increased richness of the tool positively affects the AR instruction quality.
- **H2:** Increased richness of the tool negatively affects mental effort.
- **H3:** Increased richness of the tool negatively affects the cognitive load.
- **H4:** Increased richness of the tool negatively affects the ease of use.

We conducted a between-subjects laboratory experiment to test our hypotheses, in which different participants tested one of the two AR authoring prototypes. For example, one participant group used an AR authoring prototype with a reduced feature set. The other participant group used a more extensive AR authoring prototype with an extended feature set. We choose a between-subject design to minimize possible learning effects when the same participants use both types of AR authoring tools sequentially to create instructions, as repetition could introduce bias to the collected data.

As part of our experiment, we collected qualitative and quantitative data using our AR authoring tool prototypes. The effectiveness, efficiency, and quality of the authored AR instructions were evaluated as dependent variables. Furthermore, the participants evaluated cognitive load, mental effort, and ease of use through self-reports as part of the post-experiment survey. As an independent variable, the richness of the AR authoring tool was examined using two AR authoring tool prototypes with different levels of richness.

At the beginning of each experiment session, we briefed our participants about our research objective - evaluating our AR authoring tool - as well as the experiment procedure. In addition, a pre-experiment written survey was used to collect demographic data and insights into participants' previous experiences with AR.

Before starting the experiment, we demonstrated the tool to the participants and explained its functions and how to operate it. After introducing the tool, a simple demonstration task was presented to the participants. Then, participants had to solve it autonomously to verify whether they understood how to use the tool. Following this introduction to the tool and successful completion of the demonstration task, no further information on how to use the tool was provided to the participants.

As the main task of the experiment, we requested participants to assemble a 2x2 IKEA KALLAX shelf following IKEA's paper-based instruction manual, which features eight assembly steps. We asked participants to use the AR authoring tool to create an AR assembly instruction, which means replicating the eight steps of the paper-based instruction and improving this instruction through AR visualizations. Participants were given 20 min to create their AR instruction. The experiment task is concluded when it is completed or time runs out. During the experiment, quantitative data was collected. While participants were working on the experiment task, a screen recording of their activities was captured on the tablet computer used to run the AR authoring prototype. In addition, a camera placed in the corner of the lab room was used to record video of the participant's actions. The variable effectiveness was measured based on how many of the eight steps of the IKEA paper instructions were implemented by the participants in AR using the tool. For the variable efficiency, the time required in seconds to reproduce the second step of the IKEA instructions in AR was measured. The quality variable is based on the average of two AR authoring experts' ratings for the AR instructions created by the experiment participants.

To conclude the experiment, we asked participants to complete a post-experiment survey. Among other questions, we used NASA-TLX [41] and RSME [42] questionnaires to measure perceived task load and perceived mental effort, respectively. For collecting data on participants' perceptions of the ease of use of our tool, we used the scale ease of use by Wixom and Todd [43].

We recruited 23 student participants for our laboratory experiment, ten male and thirteen female. Of these student participants, 18 are enrolled in business administration and five in industrial engineering. Although participation in the experiment was voluntary as a reward, participants received three bonus points for a written exam.

Regarding the experience with AR, 18 participants stated that they have experience with using AR to varying degrees, mainly AR games, social media filters, and shopping applications. Five participants indicated that they had never used any AR application before. Two participants stated that they'd developed AR applications before. None of the participants used an AR authoring tool before this study.

5 Results

We conducted a statistical analysis of the dependent variables collected through our laboratory experiment to test our hypotheses. Given that t-test postulates normally distributed and homogeneous variables tested our variables for normal distribution with the Shapiro-Wilk-Test [44] and for homogeneity of variance using Levene's test [45]. As shown in Table 1, the two tests show that a normal distribution and homogeneity can be assumed for all examined dependent variables, as all values are significant at the $\alpha <$ 0.05 level.

G*Power [46] was used to calculate the achieved power for an independent group's t-test for two groups of 11 and 12 participants, respectively. We chose an alpha level of $\alpha = 0.05$, and the effect size was estimated to be $d = 0.5$ based on Cohen's [47] guidelines, as there is no comparable study yet from which the effect size could be adapted. Using these values, the study achieved a power level of 0,21, which shows that the study is underpowered. Therefore, any statistical results obtained must be considered with caution.

Hypothesis H1a states that increasing tool richness has a positive effect on effectiveness, which cannot be confirmed through our t-test due to a lack of statistical significance ($t(21) = 0{,}393$, $p < 0{,}699$). However, descriptive statistics imply the opposite of our hypothesis. Higher richness slightly decreases the effectiveness. The mean effectiveness of all participants is 0.367 (SD = 0.238). The mean effectiveness of the group using the reduced tool is 0,387 (SD = 0,265), and the mean value for the more extensive tool is lower at 0,348 (SD = 0,22).

Hypothesis H1b states that increasing tool richness positively affects efficiency. This cannot be confirmed through our t-test due to a lack of statistical significance ($t(21) = -1{,}142$, $p < 0{,}266$). Our descriptive results indicate that increased richness could positively affect efficiency. The mean efficiency of all participants is 204,391 (SD = 132,608). The mean efficiency of the group using the reduced tool is 171,636 (SD = 117,150), and the mean value for the more extensive tool is higher at 234,417 (SD = 143,655).

Table 1. Calculated results for normal distribution and homogeneity.

	Shapiro Wilk Test		**Levene Test**	
Dependent Variable	Sig. (p)	Result	Sig. (p)	Result
Effectiveness	0,104	ok	0,466	ok
Efficiency	0,516	ok	0,707	ok
Quality	0,103	ok	0,993	ok
Cognitive Load	0,376	ok	0,191	ok
Mental Effort	0,013	ok	0,828	ok
Ease of Use	0,132	ok	0,626	ok

Hypothesis H1c states that increasing tool richness positively affects the quality of the AR instruction. Our t-test results show that there is a statistically significant difference between the group using the reduced prototype and the group using the extensive prototype ($t(21) = -2{,}301$, $p < 0{,}032$). Descriptive statistics also show a difference in the quality of the two groups. The mean quality score of all participants is 2,935 (SD = 1,048). The mean quality of the group using the reduced tool is 2,455 (SD = 1,011), and the mean for the extensive tool is higher at 3,375 (SD = 0,908).

Hypothesis H2 states that increasing tool richness negatively affects cognitive load, which is increased when richness is increased. This cannot be confirmed through a t-est due to a lack of statistical significance ($t(21) = -0{,}517$, $p < 0{,}61$). However, the descriptive statistics indicate a higher cognitive load when using the extensive tool. The mean cognitive load of all participants is 4,522 (SD = 1,380). The mean cognitive load of the group using the reduced tool is 4,364 (SD = 1,132), and the mean for the extensive tool is higher at 4,667 (SD = 1,611).

Hypothesis H3 states that increasing tool richness negatively affects mental effort, which is increased when richness is increased. This cannot be confirmed through our t-test due to a lack of statistical significance ($t(21) = -0{,}683$, $p < 0{,}502$). The descriptive statistics also show a higher mental load when using the extensive tool. The mean mental effort of all participants is 41,478 (SD = 30,336). The mean mental effort of the group using the reduced tool is 36,909 (SD = 30,989), and the mean value for the extensive tool is higher at 45,667 (SD = 30,455).

Hypothesis H4 states that increasing tool richness negatively affects the ease of use, which is decreased when richness is increased. This cannot be confirmed through our t-test due to a lack of statistical significance ($t(21) = 0{,}466$, $p < 0{,}646$). However, the descriptive statistics show higher ease of use of the reduced tool. The mean ease of use reported by all participants is 4,246 (SD = 1,429). The mean ease of use reported by the group using the reduced tool is 4,394(SD = 1,497), and the mean value for the extensive tool is lower at 4,111 (SD = 1,417).

6 Discussion

This paper presents an AR authoring tool that enables non-developers to create AR instructions to support complex tasks. The proposed design principles contribute theoretically to design knowledge by guiding the development of an AR authoring tool for the industrial context. The laboratory experiment results show that the design principles and their instantiation in a software prototype enable non-developers to create AR instructions self-sufficiently. Along with the existing AR literature, we found that creating AR content with complex visualization elements (i.e., various abstract AR elements, including media) can achieve better results [13, 35]. Thus, the laboratory experiment demonstrated the significantly better quality of the AR instructions without significantly increasing the participants' cognitive load and mental effort. Furthermore, there was no significant difference in the ease of use of the tools. When instating the design principles in a software prototype, we implemented passive tracking that is not perceived by the participants, intending to reduce the tool's richness [28] and thus improve the learning outcome [25]. Consequently, the AR content could not be placed precisely enough, which some participants perceived as a hindering factor. The next design cycle should investigate how active tracking affects the cognitive load and the participants' task performance. To sum up, our three theoretically grounded design principles provide prescriptive knowledge about the design of AR authoring tools for creating AR instructions. Following the DSR contribution framework of Gregor and Hevner, we consider our contribution as an improvement since we propose, on the one hand, a software prototype for a real-world problem and, on the other hand, knowledge in the form of operational design principles [27].

Although we conducted the DSR project and evaluation described in this research according to established guidelines, there are limitations that require further research. First of all, the evaluation was conducted only with participants who do not correspond to the target group (i.e., service technicians). In order for the results to be richer and more generalizable, the evaluation must be conducted with participants from the target group. Second, this study evaluates only a first step toward examining cognitive load during AR creation. Further evaluations are needed to make a validated statement about the progression of cognitive impact on users. Considering social cognitive theory and associated application self-efficacy may also raise new insights regarding cognitive effects during the AR creation process. Third, although two AR authoring experts evaluated the quality of AR instructions created by participants, a representative target group may perceive the quality of AR-based instructions differently. In a two-stage evaluation, the AR-based instructions created by the non-experts could be assessed by other non-experts, which could lead to new and representative results. Finally, due to the small sample size and limited statistical significance of the results of the t-test analysis, some caution must be exercised in regard to the generality of our findings. Nevertheless, our current statistical data, including those from the descriptive analysis, seems to confirm our formulated hypotheses. The hypotheses of this study will be tested again in an upcoming study with a larger sample size.

Acknowledgments. This research was sponsored by the German Federal Ministry for Education and Research in the project WizARd under the reference 02K18D180. Further information can be found at: https://wizard.tu-dortmund.de/

References

1. Zolnowski, A., Schmitt, A.K., Böhmann, T.: Understanding the impact of remote service technology on service business models in manufacturing: from improving after-sales services to building service ecosystems. In: 19th European Conference on Information Systems (2011)
2. Morana, S., Schacht, S., Scherp, A., Maedche, A.: A review of the nature and effects of guidance design features. Decis. Support Syst. **97**, 31–42 (2017)
3. Kaul, M., Storey, V.C., Woo, C.: A framework for managing complexity in information systems. J. Datab. Manag. **28**, 31–42 (2017)
4. Bräker, J., Osterbrink, A., Semmann, M., Wiesche, M.: User-centered requirements for augmented reality as a cognitive assistant for safety-critical services. Bus. Inf. Syst. Eng. (2022)
5. Choi, T.-M., Kumar, S., Yue, X., Chan, H.-L.: Disruptive technologies and operations management in the Industry 4.0 era and beyond. Prod. Oper. Manag. **31**, 9–31 (2022)
6. Becker-Kornstaedt, U., et al.: Support for the process engineer: the spearmint approach to software process definition and process guidance. In: 11th International Conference on Advanced Information Systems Engineering, CAiSE 1999 (2011)
7. Serván, J., Mas, F., Menéndez, J.L., Ríos, J.: Using augmented reality in AIRBUS A400M shop floor assembly work instructions. In: The 4th Manufacturing Engineering Society International Conference, vol. 1431, pp. 633–640 (2011)
8. Nebeling, M., Speicher, M.: The trouble with augmented reality/virtual reality authoring tools. In: 17th IEEE International Symposium on Mixed and Augmented Reality, pp. 333–337 (2018)
9. Palmarini, R., Erkoyuncu, J.A., Roy, R., Torabmostaedi, H.: A systematic review of augmented reality applications in maintenance. Robot. Comput.-Integr. Manufact. **49**, 215–228 (2018)
10. Konopka, B., Hönemann, K., Brandt, P., Wiesche, M.: WizARd: a no-code tool for business process guidance through the use of augmented reality. In: Demonstration & Resources Track, Best BPM Dissertation Award, and Doctoral Consortium at BPM 2022 Co-located with the 20th International Conference on Business Process Management (2022)
11. Gattullo, M., Laviola, E., Uva, A.E.: From therbligs to visual assets: a technique to convey work instructions in augmented reality technical documentation. In: International Joint Conference on Mechanics, Design Engineering and Advanced Manufacturing, pp. 1327–1339 (2022)
12. Tang, A., Owen, C., Biocca, F., Mou, W.: Comparative effectiveness of augmented reality in object assembly. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 73–80 (2003)
13. Jasche, F., Hoffmann, S., Ludwig, T., Wulf, V.: Comparison of different types of augmented reality visualizations for instructions. In: CHI Conference on Human Factors in Computing Systems (2021)
14. Kuechler, B., Vaishnavi, V.: On theory development in design science research: anatomy of a research project. Eur. J. Inf. Syst. **17**, 489–504 (2008)
15. Ashtari, N., Bunt, A., McGrenere, J., Nebeling, M., Chilana, P.K.: Creating augmented and virtual reality applications: current practices, challenges, and opportunities. In: CHI Conference on Human Factors in Computing Systems (2020)

16. MacIntyre, B., Gandy, M., Dow, S., Bolter, J.D.: DART: a toolkit for rapid design exploration of augmented reality experiences. In: 17th Annual ACM Symposium on User Interface Software and Technology, pp. 197–206 (2004)
17. Zhu, J., Ong, S.K., Nee, A.: A context-aware augmented reality assisted maintenance system. Int. J. Comput. Integr. Manuf. **28**, 213–225 (2015)
18. Krings, K., Weber, P., Jasche, F., Ludwig, T.: FADER: an authoring tool for creating augmented reality-based avatars from an end-user perspective. In: Mensch und Computer 2022, pp. 52–65 (2022)
19. Damarowsky, J., Kühnel, S.: Conceptualization and design of a workflow management system front end for augmented reality headsets. In: 30th European Conference on Information Systems (2022)
20. Porter, M.E., Heppelmann, J.E.: Why every organization needs an augmented reality strategy. Harv. Bus. Rev. 1–13 (2017)
21. Klinker, K., et al.: Structure for innovations: a use case taxonomy for smart glasses in service processes. In: Multikonferenz Wirtschaftsinformatik, pp. 1599–1610 (2018)
22. Kortekamp, S.S., Werning, S., Ickerott, I., Thomas, O.: The future of digital work - use cases for augmented reality glasses. In: 27th European Conference on Information Systems - Information Systems for a Sharing Society (2019)
23. Chen, H., Chen, C., Sun, G., Wan, B.: Augmented reality tracking registration and process visualization method for large spacecraft cable assembly. In: Augmented reality tracking registration and process visualization method for large spacecraft cable assembly, vol. 612 (2019)
24. Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol. Rev. **63**, 81–97 (1956)
25. Sweller, J.: Cognitive load during problem solving: effects on learning. Cognit. Sci. **12**, 257–285 (1988)
26. Hou, L., Wang, X., Bernold, L., Love, P.E.D.: Using animated augmented reality to cognitively guide assembly. J. Comput. Civ. Eng. **27**, 439–451 (2013)
27. Gregor, S., Hevner, A.R.: Positioning and presenting design science research for maximum impact. MIS Q. **37**, 337–355 (2013)
28. Azuma, R.T.: The most important challenge facing augmented reality. Pres.: Teleoper. Virt. Environ. **25**, 234–238 (2016)
29. Peffers, K., Rothenberger, M., Tuunanen, T., Vaezi, R.: Design science research evaluation. In: Peffers, K., Rothenberger, M., Kuechler, B. (eds.) DESRIST 2012. LNCS, vol. 7286, pp. 398–410. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-29863-9_29
30. Hönemann, K., Osterbrink, A., Wiesche, M.: Enabling non-professionals users to bring physical processes into the industrial metaverse. In: 21st Annual Pre-ICIS Workshop on HCI Research in MIS (2022)
31. Lauterbach, J., Mueller, B., Kahrau, F., Maedche, A.: Achieving effective use when digitalizing work: the role of representational complexity. MIS Q. **44**, 1023–1048 (2020)
32. vom Brocke, J., Schmiedel, T., Recker, J., Trkman, P., Mertens, W., Viaene, S.: Ten principles of good business process management. Bus. Process. Manag. J. **20**, 530–548 (2014)
33. Azuma, R.T.: A survey of augmented reality. Pres.: Teleoper. Virt. Environ. **6**, 355–385 (1997)
34. Gassen, J.B., Mendling, J., Thom, L.H., de Oliveira, J.P.M.: Towards guiding process modelers depending upon their expertise levels. In: Advanced Information Systems Engineering Workshop (2015)
35. Lavric, T., Bricard, E., Preda, M., Zaharia, T.: A low-cost AR training system for manual assembly operations. Comput. Sci. Inf. Syst. **19**, 1047–1073 (2022)
36. Subramani, M., Wagle, M., Ray, G., Gupta, A.: Capability development through just-in-time access to knowledge in document repositories: a longitudinal examination of technical problem solving. MIS Q. **45**, 1287–1308 (2021)

37. Dourish, P.: The appropriation of interactive technologies: some lessons from placeless documents. Comput. Support. Coop. Work **12**, 465–490 (2003)
38. Lieberman, H., Paternò, F., Klann, M., Wulf, V.: End-user development: an emerging paradigm end user development. Hum.-Comput. Interact. Ser. **9**, 1–8 (2006)
39. Klinker, K., Wiesche, M., Krcmar, H.: Digital transformation in health care: augmented reality for hands-free service innovation. Inf. Syst. Front. **22**, 1419–1431 (2020)
40. Park, S.-M., Kim, Y.-G.: A Metaverse: taxonomy, components, applications, and open challenges. IEEE Access **10**, 4209–4251 (2022)
41. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research human mental workload. Adv. Psychol. **52**, 139–183 (1988)
42. Zijlstra, F.R.H., van Doorn, L.: The construction of a scale to measure perceived effort (1985)
43. Wixom, B.H., Todd, P.A.: A theoretical integration of user satisfaction and technology acceptance. Inf. Syst. Res. **16**, 85–102 (2005)
44. Shapiro, S.S., Wilk, M.B.: An analysis of variance test for normality (complete samples). Biometrika **52**, 591 (1965)
45. Levene, H.: Robust tests for equality of variances. In: Olkin, I., Ed., Contributions to Probability and Statistics, pp. 278–292 (1960)
46. Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A.: G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods **39**, 175–191 (2007)
47. Cohen, J.: A power primer. Psychol. Bull. **112**, 155–159 (1992)