

Human-Centered Development and Evaluation of an AR-Assistance System to Support Maintenance and Service Operations at LNG Ship Valves

Hendrik Stern, Rieke Leder, Michael Lütjen, Michael Freitag

1. Introduction

Liquefied natural gas (LNG) is used for transportation, due to the significantly higher density in the cryogenic state compared to normal natural gas. LNG can be transported to or from regions efficiently that cannot be connected to pipeline systems, e.g., Japan or South Korea (Mokhatab 2013). In the producing regions, natural gas is liquefied, i.e., cooled down to -162°C , transported on gas tankers, initially stored at the destination port and, if required, re-vaporized and fed into the regional gas supply network (Mokhatab et al. 2006).

More complicated and not yet established is the small-scale supply of ships with LNG. Here, the cryogenic natural gas must be brought directly to the consumer in cryogenic pressure tanks by truck. This requires an adequate infrastructure and the ships must be equipped for this purpose (Mokhatab 2013). In the field of LNG transportation, special safety requirements must be observed, especially because of the risk of explosion and the extreme danger posed by cryogenic gases to living beings (freezing) (Mokhatab et al. 2006). There are currently 170 ships with LNG propulsion in service, 35 under construction, and 112 ships "LNG ready" (as of 2019, Statista 2021). At the beginning of the development, mainly ferries and supply vessels were built with such propulsion systems, which were always bunkering at fixed locations. Today, however, ships are also in operation that rely on a global supply of LNG (e.g., cruise ships) (SZ 2018). Figure 1 shows a ship-to-ship LNG bunkering process.



Figure 1: Ship-to-ship LNG bunkering process (photo: © vladsv/Fotolia.com).

To sum up, the use of LNG propulsion in ships has great environmental benefits but also creates challenges in handling the LNG. Due to the safety regulations, the maintenance of the LNG ship systems is very important and requires a high degree of reliability and accuracy. Therefore, the service technicians need a special training, which is oriented to the „IGF“-standard for the loading and unloading of liquefied gas tankers (International Maritime Organization 2017). Apart from the training, digital assistance systems, e.g., based on Augmented Reality (AR) technology, can be used to support the maintenance and service purposes of LNG ship systems. Here, assistance systems offer the possibility of virtually displaying additional information directly on the objects to be serviced. Simultaneously, the interaction with the user can be designed so that the information can be accessed according to the user's needs. In this way, the work process can be facilitated and its performance improved (e.g., avoiding errors or increasing the speed of work). Nowadays, performant end-devices are available, which enable the creation of meaningful AR-based assistance systems but the user-centric development of the user-interfaces is still challenging and decides about the real value.

Consequently, within the addressed use case of handling and transferring LNG on and between ships, reliability, and system safety can be increased due to an optimized service and maintenance process. Thus, the goal of the funded project "LNG Armaturen" was the development and evaluation of a digital assistance system using AR technology with focus on the valves. The development approach was based on the human-centered design process for interactive systems, according to DIN 9241-210. Subsequently, an evaluation was carried out through a user study and an expert review. The developed assistance system based on an Android

smartphone enables users to access maintenance instructions and manuals, simplifies spare parts' ordering, and supports the work process step by step through context-sensitive virtualizations. Its evaluation was conducted as a combined quantitative and qualitative user study in the laboratory, which included the processing of a maintenance activity at an exemplary LNG valve (Figure 2). Here, various parameters were recorded, which enabled an evaluation of the developed assistance system. Besides, an expert review with service technicians from the field contributed important findings on the usability of the developed assistance system. Overall, the prototypical assistance system offers promising potentials for reducing workload and improving processes.

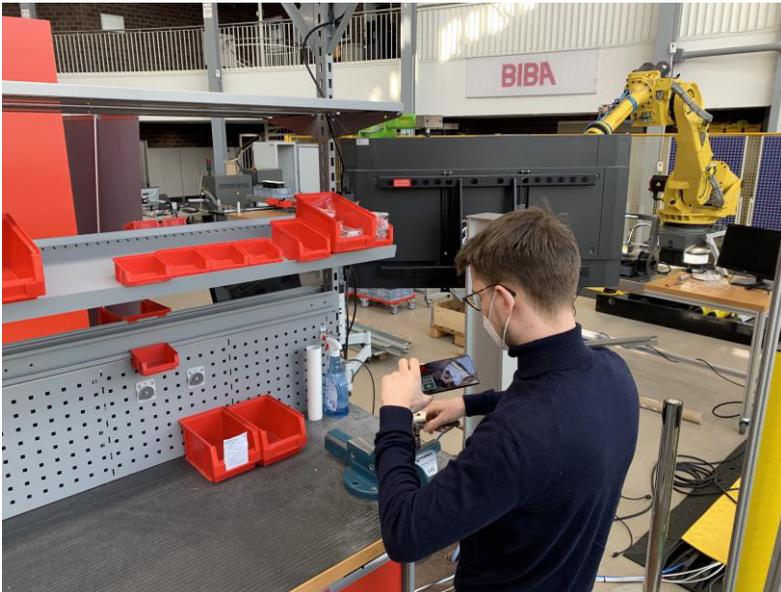


Figure 2: A test person is using the developed AR-application while working on the LNG valve during the user study.

In the following, we first put a light on existing work regarding AR assistance systems for service and maintenance, focusing on LNG use cases and AR technology (Section 2). On this basis we continue by a detailed description of our human-centered development approach with several iterations and the created AR assistance system (Section 3). Further, as being of main interest in this research work, we describe the followed approach of evaluating the developed assistance system in terms of a description and an analysis of the conducted usability study and the performed expert review (Section 4). Finally, we finish our paper by a discussion of our results and an outlook on future research work (Section 5).

2. State of Research

2.1 Industrial AR Assistance Systems

AR offers the possibility of adding virtual objects to the real world allowing users to see an enhanced version of reality (Tönnis 2010). According to Azuma (1997), an AR application must have at least the following characteristics: A combination of reality and virtuality, a real-time interaction between users and the application, and a 3D reference of virtual and real objects. AR content can be displayed via data glasses, head-mounted displays (HMD), smartphones or tablet computers (Mayer 2016). The way and quality of integrating virtual representations into reality is described by the term immersion. Immersion is disturbed, for example, by jerky images, high latencies or unrealistic representations (Butz/Krüger 2017).

Quandt et al. (2018) identified potential for AR use in industry primarily in the areas of education and training, assembly support, plant construction, predictive maintenance, quality assurance, product development, production logistics and remote maintenance. Furthermore, indoor navigation and materials management are use cases for AR applications (Fraga-Lamas et al., 2018). In contrast to paper-based information provision, AR can reduce the cognitive workload of users. This is done by virtually overlaying information depending on the context (Gattullo et al., 2020). In this way, complexity can be reduced and errors can be avoided.

When using AR, the user-oriented design of the applications plays an essential role, as this forms the interface between the users and the technology (Mayer 2016). In that way, the usability of AR applications can be increased, which in turn results in higher user acceptance (Dey et al., 2018). Other factors that influence the usability of the AR application are the used hardware (Egger/Masood, 2020), the perceived efficiency of the application from the user's point of view (Syberfeldt et al., 2016) and the cognitive load resulting from the use of the application (Egger/ Masood, 2020).

2.2 AR for Service and Maintenance Support

Several attempts to integrate AR in service and maintenance already exist. Yet, AR is not the only assisting tool in the industrial sector. Collaborative robots are also used and researched (Wolfartsberger 2019). Our interest, however, is in AR service and maintenance systems. In many cases, AR is used as a replacement for paper-based instructions or 2D references (Natakuaithung 2020, Hynes 2019, Chu 2020, Kastner 2020). 2D references lack interactivity and depth, therefore, a 3D view (with AR features) is a more helpful assistance tool (Natakuaithung 2020).

When handling complex constructions in 2D-paper-based representations, users face difficulties in understanding the drawing constructions (Chu 2020). To overcome paper-based manuals, AR-based assistance system with instant interaction can be one solution (Chu 2020, Hynes 2019). An AR instruction is of practical use,

since it leads to a lower error rate compared to a paper-based instruction though it can have a higher time consumption (Chu 2020, Kästner et. al. 2020). Other studies could prove a better time efficiency when using an AR approach (Hynes 2019, Kästner et. al. 2020). Another impact can be a lower stress level and success rate (Hynes 2019). For the creation of the AR version, the inefficiency of paper-based instructions should be defined in beforehand (Chu 2020). As tool for implementation, Unity in combination with Vuforia is often used (Chicaiza et al. 2018, Chu 2020, Hynes 2019).

Most related to our concept is the application and study by Chicaiza et al. (2018). An AR smartphone app for training and assistance in the handling of equipment and engineering was developed. In the implementation, physical manuals were transformed to a digital AR format. The AR overlay happens through graphical and text information of the scanned object. In addition, there are audio feedback, explanation videos, and the possibility to save screenshots or images. To prove the usability of their implementation, Chicaiza et al. (2018) conducted a user study using the system usability score (SUS). Their total SUS had the value of 78.43 %, which states high usability, and therefore, they could find their app suitable for training and correctly managing the use of equipment in the industrial field.

The current state of research is comparing a paper-based manual with a digital version using an AR overlay. However, most research does not consider a comparison between different variations of AR and UI. Keeping this in mind, we defined the goal of our study to compare a non-AR variation with an AR-variation of a similar UI and follow a more user-centered design approach.

2.3 Designing an Evaluating User Interfaces for AR assistance systems

Adequate user interface design has a direct impact on its usability for users. Consequently, some guidelines standardize the design of user interfaces (Shneiderman et al. 2018). Besides, step-by-step process models are available to structure the design process to avoid undesirable developments and support the development of a solution, e.g., the human-centered design process for interactive systems according to DIN EN ISO 9241-210 (DIN 9241).

Guidelines for the design of a user interface provide specific instructions, like the layout of images or text (e.g., Google 2020b) to achieve a desired final result. AR design requires to support users being immersed in the AR world and not being disturbed by a flood of information (Stern 2020). Consequently, push or full-screen notifications or a constant display of 2D content for example should be avoided. Possible user inputs should be intuitive and straightforward. The continuity of the AR experience should be in the foreground, and the user's physical interaction should be reduced. If users are already familiar with certain types of interaction (e.g., direct manipulation by dragging), these should be implemented with priority to avoid unnecessary learning efforts (Google 2020a).

User interface design guidelines and principles provide essential guidance on developing a user interface but do not evaluate its usability. For this purpose, different evaluation options are available. They aim to identify weak points of the design and to assess individual user satisfaction. For the evaluation of interactive systems, evaluation methods by experts or by users of the system are most common (Dix 2004). Usually, different expert- and user-centered evaluation methods are combined (Billingham 2014).

A user-centered evaluation reduces the risk of developing systems that do not meet users' needs. This way, prototypes can both be tested at an early stage, and the fulfillment of user requirements can be promoted. Also, qualitative feedback can support problem identification and enable immediate improvement or consideration for a later system version (Heinecke 2012).

A standard method for evaluating interactive systems are experimental evaluation studies (Jacko 2012). They can be used to show whether there are causal relations between the design of a human-machine system and effects on the humans involved, the system, or the overall system performance. Therefore, in an experimental study, one or more variables are modified in order to induce observable effects (Jacko 2012, Wickens 2014). The gain of knowledge is mainly achieved through social scientific data acquisition techniques such as interviews or questionnaires (Badke-Schaub 2012) and can be used to facilitate design decisions (Stern 2019).

In production and logistics scenarios, we usually focus on the system's usability to improve the interaction between users and technology. Therefore, questionnaires for assessing the perceived usability of an interactive system such as the System Usability Scale (Brooke 1996) can be used. The goal of a subsequent statistical evaluation of the experiments is to answer whether the selected independent variables (factors) influence the dependent variables.

3. Approach and Solution

According to the challenges presented at the beginning of Section 1 dealing with the handling of LNG, an AR solution was developed that can be used for maintenance and service purposes of LNG valves. In this context, AR offers the possibility of virtually displaying additional information directly on the valves to be maintained. This way, the user is able to call up information as needed. Thus, we aim to improve the performance of the work process (e.g., number of errors, work perception) using AR support.

We developed the AR solution according to the needs of the technicians. Here, the process for designing interactive systems according to DIN 9241-210 was used as a guideline. Its steps, such as understanding the context of use, specifying the

requirements for use, developing design solutions, and evaluating them from the user's perspective, were carried out.

Based on user stories of the employees, we initially determined the requirements in order to increase the added value of the assistance system. The result is a concept for AR-based support, which provides assistance for component replacement in the context of regular maintenance or critical malfunctions. For this purpose, we focused on the activities "retightening the valve" and "replacing the gasket". These activities are often carried out onboard the ships on which the valves are used. Consequently, it is often impossible to send a service technician, and the AR solution must be suitable for different user groups and usable without additional hardware. Therefore, an Android smartphone solution is used by providing a context-sensitive AR-based work support application.

The application was developed using Vuforia software and Unity game engine. We provided assistance through visualizations in the form of displaying the process steps that need to be carried out. The identification of the components of the LNG fitting is image-based using markers. The markers are collected in an image library, with each marker having a link to one certain valve. This provides an clear identification between them and additional markers can be added easily. In the first step, features of the markers are identified, such as edges and corners. To search for the marker in the image stream, geometric patterns that could correspond to the marker are searched and compared with those of the marker being searched for. If there is a match, size and deformation of the pattern can be used to infer the distance and viewing direction to the marker. Due to its widespread use in the AR segment, Vuforia was chosen as the framework for development (for more details on the design process, see Section 3.2). Early in the design process, mock-ups of the planned application were created and assessed, and qualitatively evaluated in collaboration with the users of the application. Improvements resulting from this were directly incorporated into the further design process (Section 3.1).

After the development of the application, an evaluation study consisting of a user test in the laboratory and an expert review in practice was conducted (Section 4).

3.1 Development of Mock-Ups

Before the actual assistance system was implemented, we developed several interactive mock-ups and iterated them for the design and its functionality. In the following, the development of the mock-up is described in detail. The tool Adobe XD was used to specify the UI based on the user requirements since it allows to share the developed UI with others.

Another important point was the consideration of valve detection. We came up with two possibilities. The first was of a simple image/tag recognition by adding tags next to the valves. The second was the valve serving as a model for detecting itself (object recognition). Here, tag detection was found to be the better solution,

due to two reasons. One consideration was the problem with the detection of the valve itself, as there are too many very similar models, with an almost impossible differentiation between them. The other consideration was its installation aboard a ship, where most valves and armatures are covered due to isolation, which also makes object recognition almost impossible.

A paper-based instruction served as the model for the creation of the mock-up. The instruction consisted of information like a sectional drawing, information about the usage of the valve (e.g., the temperature ranges the valve can be used in) and step-by-step instructions for several service and maintenance tasks. For the processes of replacing and tightening we used the already existing picture instructions divided into several mounting steps. Figure 3 shows the instruction for the first mounting step for replacing the gasket.

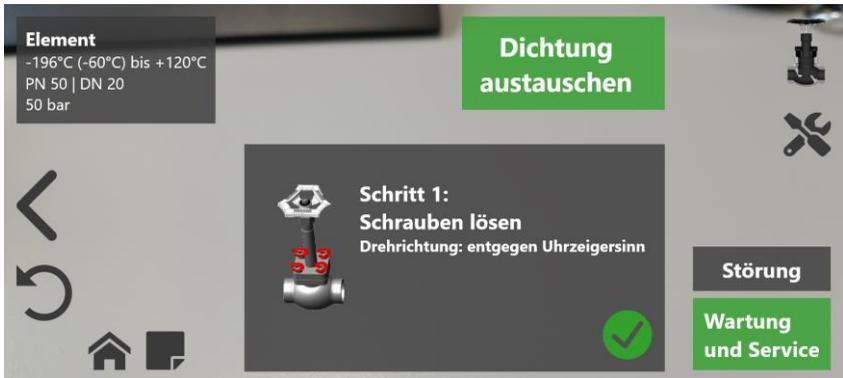


Figure 3: Screenshot from the mock-up with the first assembly step of the task “replacing the gasket” as an example. It displays information on the detected valve (left), on the current step of the task (center), and on the navigation within the app (right).

With relation to the paper manual, we designed the interface for the application. After the valve was detected, we thought about the possibility to define the casing of the object, which can cause a change in the torque of some screws. This step is followed by the main menu. Here and for all following pages, we added a back, rescan, notes, valve, and home button. When coming to the actual maintenance and service tasks, we thought about two ways leading to the instruction steps. On one hand, we designed a navigation when only the problem is known, but not its remedy. On the other hand, we offer a navigation directly to the remedy. It leads the user directly to the instructions, in our case “tighten” and “replace the gasket”. After the remedy was chosen, the instruction steps are displayed with internal navigation of going forward and backward only affecting the maintenance steps. In addition, we decided on an extra button in the maintenance steps for the needed tools to complete the task.

By linking the created pages within Adobe XD, the mock-up became interactive with the buttons leading to the page of its purpose. One example of navigation is shown in Figure 4.

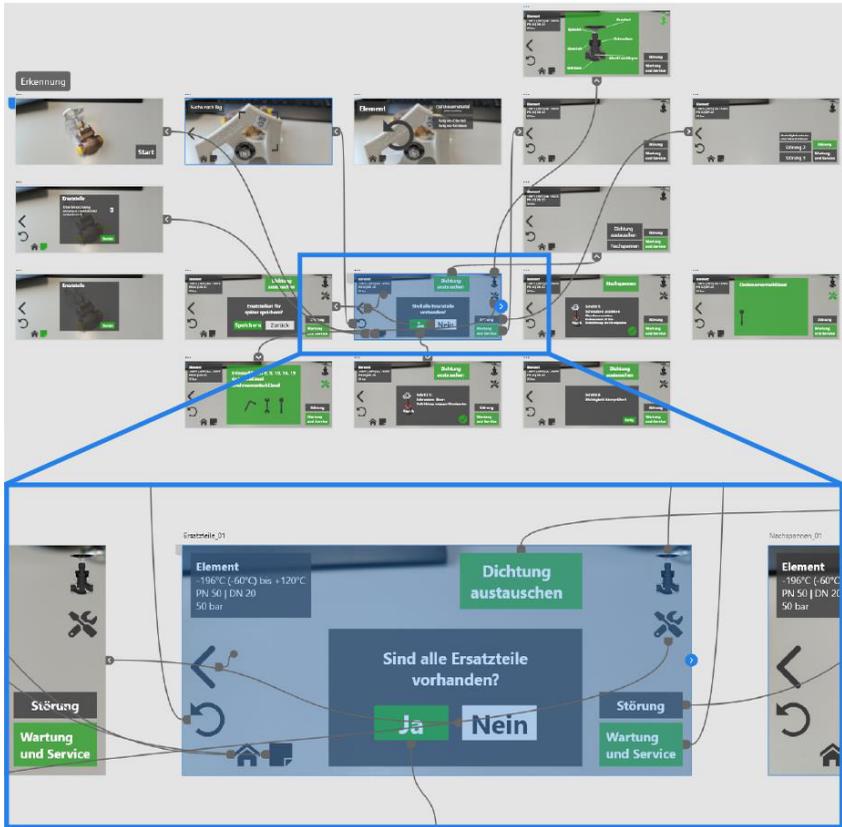


Figure 4: Example of the interactions between different pages (curved grey lines) in the mock-up when interacting with a button.

A finished first interactive mock-up was discussed with two expert users. The main discussion points were the position of the tag, a need for a direct help button, the side of main menu buttons, an overview about spare parts, the torque value, and another maintenance step.

The tag was thought to fit best on the handwheel of a valve, as it would be too small and hard to replace. However, a final solution for the placement was not found. Whereas the main menu could be arranged anew, by switching the side, following the reading direction. A direct help button was also added, for a better

communication to the service as well as the possibility to save and send an order request of spare parts.

An important point for replacing the gasket maintenance steps, is to inform the user about depressurizing and rendering inert the installation as mandatory, before disassembling the valve. Additionally, a button for the required tools was found to be useful. Bearing all this feedback in mind, we continued to the implementation of the application. Now, the mock-up served as a guideline in the final development phase.

3.2 Development of the AR-Assistance System

The application was finally implemented in Unity. According to the feedback for the mock-ups, the menu was switched to the left side and a direct help button with links to the partner website, a service mail address, a phone number as well as the pre-step (step 0) were considered.

The AR function was added with the help of the add-on Vuforia. Vuforia provided a program in which 3D models can be imported and used to train a machine learning algorithm for object recognition. This model can be imported into Unity and used together with the Vuforia add-on. Since the object recognition happens directly at the valve, the UI had to be re-thought. Now, the screen should have more space for the camera and the information had to be placed in a smaller text field. With several layout attempts, the right proportions were found. Figure 5 shows an example for the AR interaction with an example step (remove top part and gasket).



Figure 5: Screenshot of the AR version of the application with the third step an example of the digital instruction.

After all previously defined functions and navigations were implemented, we considered a second version of the application with less AR features and tag detection. Therefore, we duplicated the Unity project and replaced the object recognition with image recognition. For this feature, we added the add-ons AR Foundation

and AR Core. The AR Foundation provides methods for image recognition. Only the image, which should be scanned, had to be added to a predefined image library. Also, we had to define what happens after the image was successfully scanned. In addition, AR Core was required to build the application for Android, otherwise, the AR functions cannot be used. As the simple instruction images have no depth, we decided to use the 3D model of the valve and added some explaining animations, showing what has to be done. An example of the implementation is shown in Figure 6.



Figure 6: Screenshot from the normal version of the application with the third step as an example of the digital instruction.

For both applications, we added a script to track the time (for each step, step time in total and overall application usage time), taps (for each step, step taps in total, and taps for the entire usage of the application), and button presses for additional comparison factors for the evaluation of both versions.

4. Evaluation

After completing the development phase of the AR application and the altered version without using AR technology, we carried out a comprehensive evaluation. As described in Section 2, we followed an approach that combines a user study in the laboratory with an expert review (Billinghamurst 2014). The underlying idea was that test users perform a specific maintenance and service task on a real LNG valve (replacement of a gasket). The test users were supported either by means of the AR app or the non-AR app. In some cases, the test users were experts who work professionally in maintenance and servicing of LNG valves. In these cases, we focused on a qualitative evaluation of the usability of the AR application in practice. During and after the execution of the task, we were able to collect information on work performance and work perception by tracking qualitative parameters and by means of a questionnaire. For this purpose, the System Usability Scale (SUS)

(Brooke 1996) and the NASA-TLX methodology (Xiao 2005) were used in the questionnaire. A combination of these two commonly used techniques (Sauro 2011, Xiao 2015) was used in our study in order to both address the usability of the applications and the cognitive and physical workload which is caused by its usage. In this section, we will go into detail about the setup and implementation of the user study and the expert review, as well as describe and evaluate the results acquired.

4.1 User Study – Laboratory Test

4.1.1 Setup

The setup of the laboratory test consisted of an assembly workstation, the LNG valve, the necessary tools, a replacement gasket and the smartphone with the application for support (Figure 7). The test sequence followed a consistent procedure in each case: First, the test administrators gave a short introduction; during the entire experiment, the test users were allowed to ask questions. Before starting the task, the test users were instructed about the specific task (replacement of the gasket). After completing the task, all test users filled out the questionnaire and were then able to give further comments on the application and the experiment in a short interview. In beforehand of the actual user study, a test run was performed.

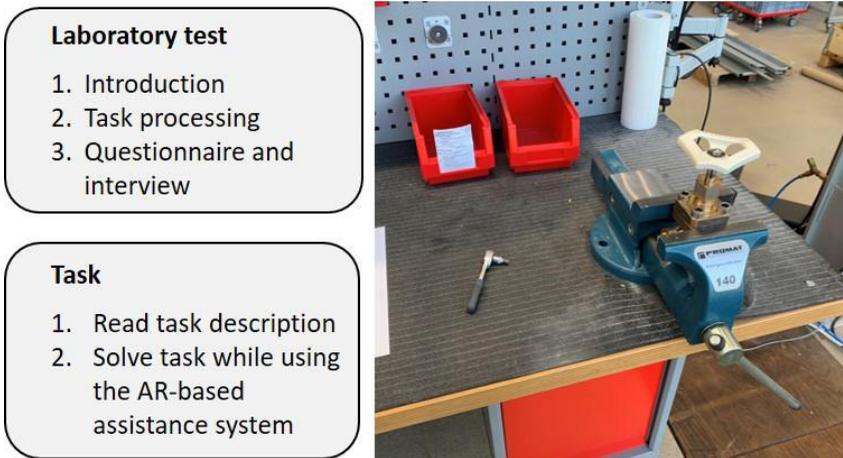


Figure 7: The setup for the laboratory user study.

We could acquire in total ten participants. Unfortunately, one participant could not speak German, which lead to a much longer finishing time and comprehension problems. For another participant, the log was not recorded correctly. Due to this, we decided not to consider them in our evaluation and results. All other participants have been male, are employed, and are between 28 and 51 years old

(mean=34.6). They stated that no one has prior experience in professional maintenance, whereas almost everyone, except for one participant, has prior experience with AR, with no especially familiarity with AR (mean=2.429, with a rating between 0 and 5).

The experimental design was that half of the test users would use the application in the AR version and the other half in the non-AR version. The allocation was randomized. As already shown in Section 3, the two applications differ in whether they use AR technology or 3D animations to illustrate the work steps. This procedure allowed us to compare the effects on work performance and work perception triggered by the AR application with those of a similar application. Thus, distortions of the results based on confounding variables (Jacko 2012), such as hardware used when comparing a paper version with an AR application, can be avoided.

The questionnaire was divided into four sections. The first section contained general demographic questions. In the second part, we asked for the system usability, using the established system usability scale (SUS) questions and the corresponding 5-point Likert scale. The SUS contains questions about the usage of the application, its complexity, and its functionality. The third section was about the task load divided into mental demand, physical demand, temporal demand, performance, effort, and frustration, following the NASA-TLX. To not confuse the participants with different point scales, we decided to choose a 5-point scale, as we used for the SUS. In the last part, we asked about the general impression of the experiment and the application with a last free to fill in a text field.

4.1.2 Results

The results can be divided into the system usability, task load, time and tap measurements, and the interview, which are described in the following.

As Figure 8 shows, the AR application has slightly lower usability. The normal application has a total mean scale of 83.75 % (SD = 4,1), while the AR version has a total mean scale of 83,125 % (SD = 9,6) which is a little below the total average of 83,4375 % (SD = 7,39). The worst achievable score is 25% and the highest is 100%.

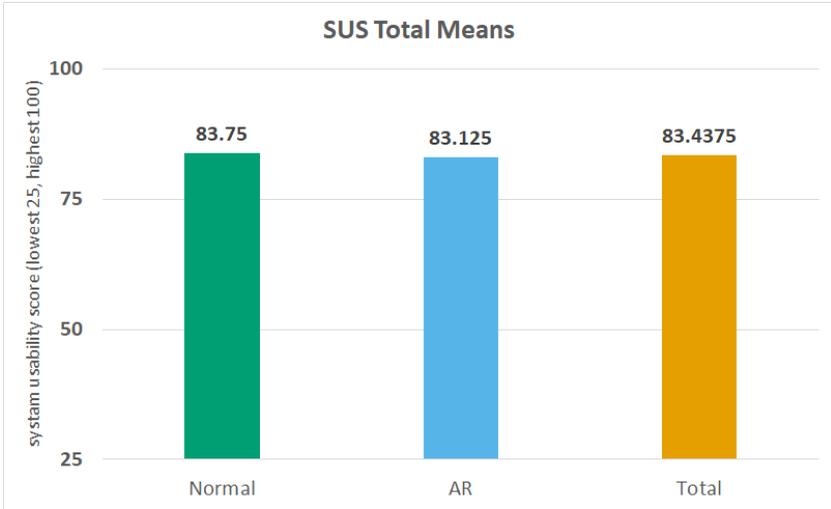


Figure 8: The system usability score means for both versions and the total average.

With a mean scale of 11.25 (SD = 1.79) points, the participants stated, that they had a little higher task load with the AR application (mean = 10.78, SD = 1.48) and the total average (mean = 11, SD = 1.66). The lowest task load possible is 6 points and the highest is 30 points, which can be seen in Figure 9.

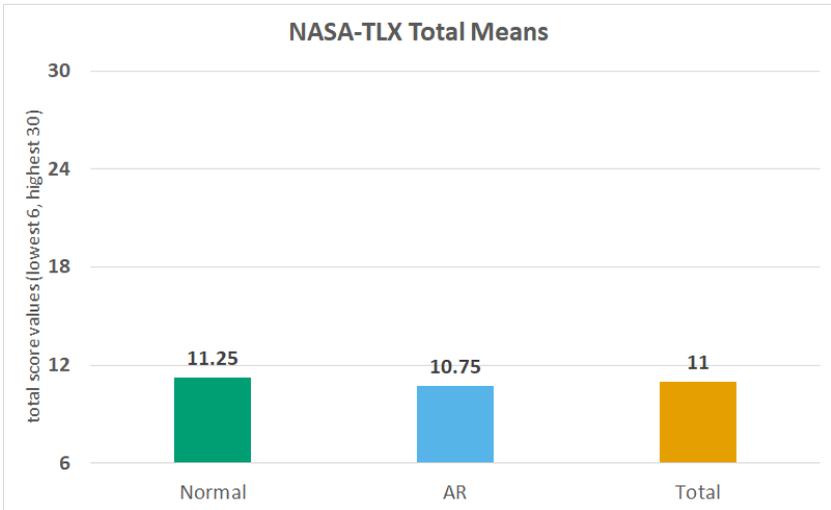


Figure 9: The NASA-TLX mean scores for both versions and the total average.

Concerning the log files, the results show that participants, using the AR application, needed almost three minutes more than compared to the normal version. With the normal version, the participants needed 04:20 minutes (SD = 00:01) and with the AR version, the participants needed 07:30 minutes (SD = 00:05). The time differences are shown in Figure 10.

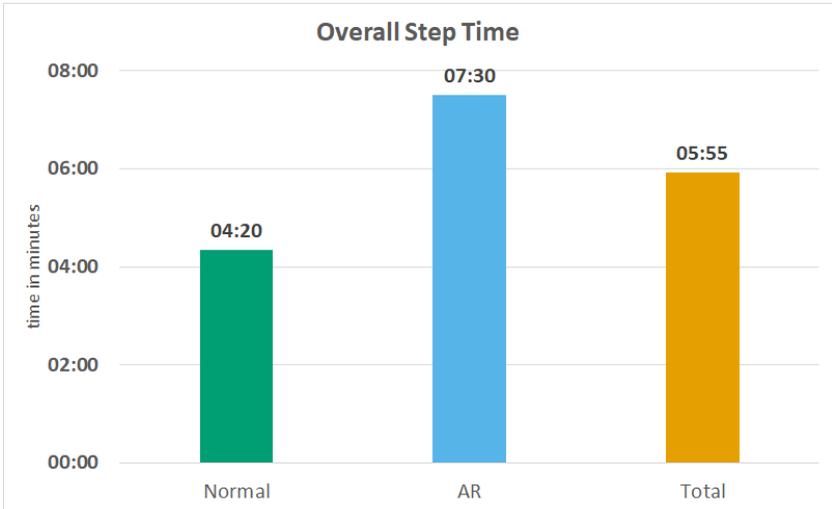


Figure 10: The means of the time needed to complete all instruction steps for both versions and the total average.

Furthermore, the participants also tapped more often on the phone screen, using the AR application, than using the normal application. With a total mean of 23 taps in AR (SD = 5.34) and 12.75 taps using the normal version (SD = 0.83), the difference is more than 10 taps, which is shown in Figure 11.

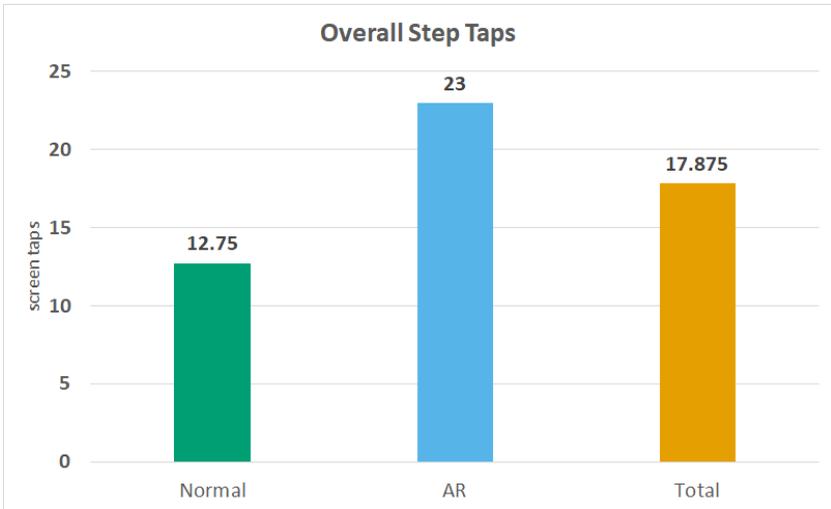


Figure 11: The mean of the taps needed to complete all instruction steps for both versions and the total average.

Besides the time and tap measurements, the logs reveal, that one participant toggled between step 1 to step 0 for five times, before continuing to step 2. Another participant first opened the steps to tighten the screws. The same participants also toggled two times between step 1 and step 2, before continuing with step 3.

During the study, we observed the participants whether all steps were followed correctly. With these observations we could figure out three types of errors. The first type was not replacing the gasket correctly. Only one participant did not replace the gasket, but adding the new one on top. The second type of error was not following the instruction at step 0 to loose the handwheel, which was not followed by six participants. The last error is not tightening the screws crosswise, not followed also by six participants. When comparing both application versions, the participants using AR (mean = 1, SD = 0.71) made half an error less than using the normal version (mean = 1.5, SD = 0.5).

After the experiment, we gave the participants the possibility to tell us about their impression of the application. The statements of the interview, combined with our observations during the task are summarized in the following.

Overall, both applications were well received, thought to be useful, especially for inexperienced people, and self-explaining. The UI was found to be good. However, some participants stated that some instruction steps are trivial and therefore, should be combined with others.

The task itself was mostly well apprehended, but some wished for an explanation video as an introduction. Only a few participants were confused when coming to the spare parts question. They could either not find the spare gasket or were expecting more than one gummy gasket, reasoned by too much information in spare part dialogue. Also, for one participant, the ending of the experiment was a little unclear. In general, that the task was found to be too simple, especially for the AR version.

The participants used the phone differently. Three participants, all using the normal version, were holding the phone in their hands only for scanning the tag and afterward placing it beneath the jaw vise. Three participants, one of them using the normal version, operated similarly but picked up the phone from time to time. The other two participants (AR) hold the phone almost permanently in their hands.

During the experiment, some participants struggled with the valve itself. In some cases, the top part was hard to remove, in others, the screws were stuck to the casing and were hard to loosen. Also, in some cases, the gasket got stuck at the top part of the casing.

The participants, using the normal application, stated that they were confused about the tag and wished for a QR code. In addition, two participants wanted different feedback after the tag was detected and wanted the tag to stay detected when using the back buttons until reaching the main menu.

On the other side, the participants using the AR version were looking for a tag instead of scanning the object itself for detection. Yet, the AR features were well received, though not all animations were found to be useful. Furthermore, when the top part was dismounted, the participants tried to scan the top part and for this, the animation was rotated for 90°. The participants also wished for a hint, when AR is really necessary, for knowing when to pick up the phone or a mix of AR and instruction videos were liked.

Other improvements were proposed by the participants, containing a highlight or confirmation (checklist) of the most important steps and lighter animations, a voice control, a bigger information field, a preview of the next step, documentation tools (take a picture, adding own text to the notes) and an English version.

Lastly, some participants mentioned, they have tightened the screws crosswise, due to previous knowledge and one participant explained that the orientation of the valve's top part was not clear when remounting it.

4.2 User Study - Expert Review

In contrast to the user study in the laboratory, we modified the procedure for the expert review due to the professional background of the two participants. Here, we chose not to provide an explicit task, but rather to support a free testing of the app. Participants were first asked to familiarize themselves with the app and then

perform the same task as in the lab test (replacing the gasket). After completion of the task, the questionnaire was used to evaluate the usability of the application as well, and the impressions of the test users were queried in a subsequent interview. Due to the different focus of the study, only the AR version of the application was used in the expert review.

During the review, we could figure out some problems regarding the application. Like in the user study, the navigation was a bit unclear at the beginning, as one expert tried to press the valve-button in order to open the disturbances. Besides, we had to point one of the experts to make use of the AR functions, which then were well received. After the troubles in the beginning were overcome, the application was used as intended, though from time to time the animation froze.

In the interview, the experts could give more information about the assembly steps. The experts mentioned there should be more emphasis on the step for loosening the handwheel. Furthermore, it was stated that the replacement or greasing of the screws with a premium steel valve should be added and highlighted and before reassembling the top part of the valve, the sealing surface should be checked for dirt and damages.

Overall, the application was well received. Especially for the maintenance on board, the experts stated a high potential of the application. It was thought to improve the service with professional instructions always available on every smartphone. However, it was also thought about the bad conditions on board (light, isolation), suggesting a combination of AR and video tutorials or even the usage of AR glasses.

5 Discussion

With both applications having a usability of more than 80 %, each version is of a high usability. This result is based on eight, non-specialist participants, which need to be taken into consideration here. Nonetheless, the participants with the normal application were faster in solving the task, but have done more mistakes and had a slightly higher task load. With relation to the use case and the sensitive safety regulations this is an important aspect. On the one hand, the time difference was caused by two participants (AR), who swapped between some steps several times. But, when calculating anew the completion times, leaving out the for and back taps, the AR version still is in average more than two minutes (mean = 06:49 instead of 07:30) slower than the normal version (mean = 04:20). On the other hand, when using the AR version, the participants have to pick up the phone more often and re-scan the objects. This will also cause a rise in time. A possible solution for this is using a tripod, so the phone stays in place and the user must not always pick up the phone and rescan the object, so the user has the hands free for the task. Alternatively, an AR-HMD can be used.

The second proposal for the task completion can also explain the difference in the phone usage. As the participants always have to pick up the phone to rescan the object for showing the animations, when using the AR version. In contrast, in the normal version, the participants could access the animation at all time and just had to tap forward or backward. With this background, the normal application was more likely for hands-free tasks.

A solution for the tracking problems in AR, when the top part of the valve is dismantled can be training an own model for the single parts. At the moment, only the entire valve exists as a model reference. Together with the predefined tracking settings, it is limited and the animation might appear rotated, as the last position of the object was saved in the memory. Another option is the possibility to overlay the AR version with instruction videos or animations or to combine the previous step with the one when the gasket is replaced, like the participants wished for.

In the expert review, the experts mentioned that the animation froze a few times, which might be due to the different phone, which was used. The object recognition and its animation are complex and as the phone used in the user study was a high-end model, the complexity was not of a problem, but on other phones, with a different processor, the performance can drop. In order to the impossibility of using the normal app, a different kind of base of an image recognition, which does not require AR Core, can be used. Then it should also be possible to install the applications on both, Android and Apple smartphones.

Finally, the task itself should be considered. Most participants stated it to be too simple, which was due to the paper-based instructions we received from the manufacturing company of the valve. At the expert review, the experts stated that the functionality would also fit for harder tasks, but for the usage on ships and several different valves, also the task we chose, was found to be reasonable. Considering the learning curve, simple task can be usefully supported with AR overlay instructions, but with some experience, a simpler instruction would probably do it. On the contrast, when the task itself is more complex, AR might reveal its full potential since it was found to be very beneficial at lowering the mental workload at complex and varying tasks, i.e. we expect a more significant difference in error rates here.

6. Summary and Conclusion

6.1 Summary

In this paper, we proposed an assistance system using AR and image recognition to support the maintenance on board of an LNG driven vessel. In order to transfer paper-based instructions for service and maintenance into an AR Assistance System, we first did a requirements analysis and translated the results into a mock-up. After professional feedback, where we received improvement suggestions for the application, we developed two different versions of the application, an AR version and a more basic approach with image recognition (tag). For their implementation, the mock-up served as model.

With the two versions of the application, we conducted a between-subjects user study. Its main goal was to compare both versions to find out which is performing better in terms of usability, task load, task completion and the error rate. Both versions showed a high usability, with the normal version having a lower task completion time, but the AR version having a lower task load and a lower error rate. Keeping in mind the difficult surrounding conditions on a ship, with mostly isolated valves, other components, and difficult lighting conditions, we conclude that a mixture of both versions and the possibility to switch off AR would fit the purpose of this application best.

6.2 Future Work

As future research work, we consider several starting points and actions. First, some changes and improvements of the AR-based assistance systems could be implemented, which are both hardware- and software-related. Here, an AR-HMD (head mounted device, e.g., Microsoft HoloLens 2) could be used instead of an AR-Handheld to improve the systems ergonomics (hands-free use). Besides, we expect a particular UI-design focusing on AR to increase the usability and performance of an AR-based assistance system. Further, as highlighted by the study participants, a mixed version of the AR-application and the normal application could combine the advantages of both systems. Here, e.g., allowing users to toggle between AR-virtualizations at complex process steps and a basic, non-AR-process guide could be helpful.

With regard to the evaluation of the AR assistance system, further user studies would be useful. Here, a larger number of test users as well as a better fit of the test users to the intended system users (e.g., with regard to prior knowledge, technology affinity) would be helpful to increase the validity of the study results. Furthermore, other, more complex tasks or components could increase the generalizability of the results.

Acknowledgments

The authors would like to thank the German Federal Ministry of Economic Affairs and Energy (BMWi) for their support within the project "LNG Transfer - LNG Armaturen Set" (grant number 16KN062728).

References

- Azuma, R. (1997): A Survey of Augmented Reality, S. 355–385.
- Badke-Schaub, P., Hofinger, G., & Lauche, K. (2012). Human Factors. Human Factors. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Billinghurst, M.; Clark, A.; Lee, G. (2014): A Survey of Augmented Reality. In: Foundations and Trends in Human-Computer Interaction 8 (2-3), S. 73–272.
- Brooke, J. (1996): SUS: A quick and dirty usability scale. In: P. W. Jordan, B. Thomas, I. L. McClelland and B. Weerdmeester (Ed.): Usability Evaluation in Industry. London: Taylor & Francis, p. 189–194.
- Butz, A.; Krüger, A. (2017): Mensch-Maschine-Interaktion. 2., extended edition. Berlin, Boston: De Gruyter Oldenbourg.
- Chicaiza, E.A., De la Cruz, E. I., Andaluz, V.H., 2018. Augmented reality system for training and assistance in the management of industrial equipment and instruments, in: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Turek, M., Ramalingam, S., Xu, K., Lin, S., Alsallakh, B., Yang, J., Cuervo, E., Ventura, J. (Eds.), Advances in Visual Computing, Springer International Publishing, Cham. pp. 675–686.
- Chu, C.H., Liao, C.J., Lin, S.C., 2020. Comparing augmented reality-assisted assembly functions - a case study on dougong structure. Applied Sciences 10, 3383. doi: 10.3390/app10103383.
- Hynes, E., Flynn, R., Lee, B., Murray, N., 2019. A quality of experience evaluation comparing augmented reality and paper based instruction for complex task assistance, in: 2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP), pp. 1–6. doi: 10.1109/MMSP.2019.8901705.
- Dey, A., Billinghurst, M., Lindeman, R. W., and Swan, J. E. (2018). A Systematic Review of 10 Years of Augmented Reality Usability Studies: 2005 to 2014. Front. Robot. AI 5, 1–28. doi: 10.3389/frobt.2018.00037.
- DIN EN ISO 9241-210:2011-01, Ergonomie der Mensch-System-Interaktion - Teil 210: Prozess zur Gestaltung gebrauchstauglicher interaktiver Systeme (ISO 9241-210:2010).
- Dix, A.; Finlay, J.; Abowd, G.; Beale, R. (2004): Human-computer interaction, 3. Auflage. Harlow: Pearson Education.

- Egger, J., and Masood, T. (2020). Augmented reality in support of intelligent manufacturing – A systematic literature review. *Computers & Industrial Engineering* 140, 106195. doi: 10.1016/j.cie.2019.106195.
- Fraga-Lamas, P., Fernandez-Carames, T. M., Blanco-Novoa, O., and Vilar-Montesinos, M. A. (2018). A Review on Industrial Augmented Reality Systems for the Industry 4.0 Shipyard. *IEEE Access* 6, 13358–13375. doi: 10.1109/ACCESS.2018.2808326.
- Google (2020a): Augmented Reality Design Guidelines. Interaction - UI. Google Inc. Available online: <https://designguidelines.withgoogle.com/ar-design/interaction/ui.html>, last accessed on 31.03.2021.
- Google (2020b): Components. Available online: <https://material.io/components>, last accessed on 31.03.2021.
- Heinecke, A. (2012): *Mensch-Computer-Interaktion*. Berlin, Heidelberg: Springer.
- International Maritime Organization (2017): Section A-V/3 Mandatory minimum requirements for the training and qualification of masters, officers, ratings and other personnel on ships subject to the IGF Code. In *STCW Code*.
- Jacko, J. A., Yi, J. S., Sainfort, F., & McClellan, M. (2012). Human factors and ergonomic methods. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (pp. 289–329). John Wiley & Sons.
- 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, 18–33. doi: 10.1016/j.chb.2018.04.054.
- Kästner, L., Eversberg, L., Mursa, M., Lambrecht, J., 2020. Integrative object and pose to task detection for an augmented-reality-based human assistance system using neural networks. URL: <https://www.researchgate.net/publication/344012887IntegrativeObjectandPosetoTaskDetectionforanAugmented-Reality-basedHumanAssistanceSystemusingNeuralNetworks>.
- Mayer, F.; Pantförder, D. (2016): Unterstützung des Menschen in Cyber-Physical Production Systems. In: Vogel-Heuser, B.; Bauernhansl, T. and ten Hompel, M. (Ed.): *Handbuch Industrie 4.0 Bd.2. Automatisierung*. 2. Auflage. Berlin, Heidelberg: Springer, S. 525–535.
- Mokhatab, S., Economides, M.: "Onshore LNG Production Process Selection." Paper presented at the SPE Annual Technical Conference and Exhibition, San Antonio, Texas, USA, September 2006. doi: <https://doi.org/10.2118/102160-MS>.
- Mokhatab, S., Mak, J. Y., Valappil, J. V., & Wood, D. A. (2013). *Handbook of liquefied natural gas*. Gulf Professional Publishing.
- Natakuaithung, P. and Kusakunniran, W., "Development of AR Learning Assistance Tool for Clay-Sculpting 3D Model," 2020 12th International Conference on Knowledge and Smart Technology (KST), 2020, pp. 109-114, doi: 10.1109/KST48564.2020.9059484.
- Quandt, M.; Knoke, B.; Gorldt, C.; Freitag, M.; Thoben, K. (2018): General Requirements for Industrial Augmented Reality Applications. In: *Procedia CIRP* 72, S. 1130–1135.
- Rapaccini, M., Porcelli, I., Espíndola, D. B., and Pereira, C. E. (2014). Evaluating the use of mobile collaborative augmented reality within field service networks: the case of Océ Italia – Canon Group. *Production & Manufacturing Research* 2, 738–755. doi: 10.1080/21693277.2014.943430.

- Sauro, J. and Lewis, J.R. (2011): When designing usability questionnaires, does it hurt to be positive? In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11). Association for Computing Machinery, New York, NY, USA, 2215–2224. DOI: <https://doi.org/10.1145/1978942.1979266>.
- Shneiderman, B.; Plaisant, C.; Cohen, M.; Jacobs, S.; Elmqvist, N. (2018): Designing the user interface. Strategies for effective human-computer interaction. Sixth edition, global edition. Boston: Pearson.
- Statista (2021): Bestand an fahrenden und bestellten Schiffen mit LNG-Antrieb sowie LNG ready Schiffen weltweit in den Jahren 2009 bis 2026, Available online: <https://de.statista.com/statistik/daten/studie/1079540/umfrage/bestand-an-lng-betriebenen-und-bestellten-sowie-lng-ready-schiffen-weltweit/>, last accessed on 23.04.2021.
- Stern, H. and Becker, T. (2019): Concept and Evaluation of a Method for the Integration of Human Factors into Human-Oriented Work Design in Cyber-Physical Production Systems. Sustainability. 11 (4508).
- Stern, H.; Quandt, M.; kleine Kamphake, J.; Beinke, T. and Freitag, M. (2020): User Interface Design für Augmented Reality-basierte Assistenzsysteme. In: Michael Freitag (Ed.), Mensch-Technik-Interaktion in der digitalisierten Arbeitswelt, Tagungsband der Wissenschaftlichen Gesellschaft für Arbeit- und Betriebsorganisation. Berlin: GITO.
- Sueddeutsche (2021): Neues Kreuzfahrtschiff Aida Nova. Available online: <https://www.sueddeutsche.de/reise/kreuzfahrt-kreuzfahrtschiff-aida-nova-1.4261903>, last accessed on 23.04.2021.
- 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, 108–113. doi: 10.1016/j.procir.2016.02.017.
- Tönnis, M. (2010): *Augmented Reality*. Berlin, Heidelberg: Springer.
- Wickens, C., Gordon, S. and Liu, Y.: (2014). „An introduction to human factors engineering“. New York, Longman.
- Wolfartsberger, J., Hallewell, Haslwanter, J., Lindorfer, R., 2019. Perspectives on assistive systems for manual assembly tasks in industry. *Technologies* 7, 12. doi: 10.3390/technologies7010012.
- Xiao, Y.-m, W. Z.-m., Wang, M.-z. and Lan. Y.-j, (2005): [The appraisal of reliability and validity of subjective workload assessment technique and NASA-task load index]. *Chinese journal of industrial hygiene and occupational diseases*, 23. Pp. 178-81.